A spine approach to branching diffusions with applications to $\mathcal{L}^p$-convergence of martingales

Robert Hardy and Simon C. Harris

Department of Mathematical Sciences, University of Bath,
Claverton Down, Bath, BA2 7AY, UK.
email: S.C.Harris@bath.ac.uk

Summary. We present a modified formalization of the 'spine' change of measure approach for branching diffusions in the spirit of those found in Kyprianou [40] and Lyons *et al.* [44, 43, 41]. We use our formulation to interpret certain 'Gibbs-Boltzmann' weightings of particles and use this to give an intuitive proof of a general 'Many-to-One' result which enables expectations of sums over particles in the branching diffusion to be calculated purely in terms of an expectation of one 'spine' particle. We also exemplify spine proofs of the $\mathcal{L}^p$-convergence ($p \geq 1$) of some key 'additive' martingales for three distinct models of branching diffusions, including new results for a multi-type branching Brownian motion and discussion of left-most particle speeds.

1 Introduction

Consider a branching Brownian motion (BBM) with constant branching rate r and offspring distribution A, which is a branching process where particles diffuse independently according to a standard Brownian motion and at any moment undergo fission at a rate r to be replaced by a random number of offspring, $1 + A$, where A is an independent random variable with distribution

$$P(A = i) = p_i, \qquad i \in \{0, 1, \ldots\},$$

such that $m := P(A) = \sum_{i=0}^{\infty} i\, p_i < \infty$. Offspring move off from their parent's point of fission, and continue to evolve independently as above, and so on.

Let the configuration of this BBM at time t be given by the $\mathbb{R}$-valued point process $\mathbb{X}_t := \{X_u(t) : u \in N_t\}$ where N_t is the set of individuals alive at time t. Let the probabilities for this process be $\{P^x : x \in \mathbb{R}\}$, where P^x is the law starting from a single particle at position x, and let $(\mathcal{F}_t)_{t \geq 0}$ be the natural filtration. It is well known that for any $\lambda \in \mathbb{R}$,

$$Z_\lambda(t) := \sum_{u \in N_t} e^{-rmt} e^{\lambda X_u(t) - \frac{1}{2}\lambda^2 t} = \sum_{u \in N_t} e^{\lambda X_u(t) - E_\lambda t} \tag{1}$$

where $E_\lambda := -\lambda c_\lambda := \frac{1}{2}\lambda^2 + rm$, defines a *positive* martingale, so $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ exists and is finite almost surely under each P^x. See Neveu [46], for example.

One of the central elements of the spine approach is to interpret the behaviour of a branching process under a certain change of measure. Chauvin and Rouault [9] showed that changing measure for BBM with the Z_λ martingale leads to the following 'spine' construction:

Theorem 1.1 *If we define the measure* $\mathbb{Q}_\lambda^x$ *via*

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^x}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = e^{-\lambda x} Z_\lambda(t), \tag{2}$$

then under $\mathbb{Q}_\lambda^x$ *the point process* $\mathbb{X}_t$ *can be constructed as follows:*

- *starting from position* x*, the original ancestor diffuses according to a Brownian motion on* $\mathbb{R}$ *with drift* λ*;*
- *at an accelerated rate* $(1+m)r$ *the particle undergoes fission producing* $1+\tilde{A}$ *particles, where the distribution of* $\tilde{A}$ *is independent of the past motion but is* size-biased*:*
$$\mathbb{Q}_\lambda(\tilde{A} = i) = \frac{(i+1)p_i}{m+1}, \qquad i \in \{0, 1, \ldots\}.$$
- *with equal probability, one of these offspring particles is selected;*
- *this chosen particle repeats stochastically the behaviour of the parent with the size-biased offspring distribution;*
- *each other particle initiates, from its birth position, an independent copy of a* $P^{\cdot}$ *branching Brownian motion with branching rate* r *and offspring distribution given by* A *(which is without the size-biasing).*

The chosen line of descent in such pathwise constructions of the measure, here $\mathbb{Q}_\lambda$, has come to be known as the *spine* as it can be thought of as the backbone of the branching process $\mathbb{X}_t$ from which all particles are born. The phenomena of size-biasing along the spine is a common feature of such measure changes when random offspring distributions are present.

Although Chauvin and Rouault's work on the measure change continued in a paper co-authored with Wakolbinger [10], where the new measure is interpreted as the result of building a conditioned tree using the concepts of Palm measures, it wasn't until the so-called 'conceptual proofs' of Lyons, Kurtz, Peres and Pemantle published around 1995 ([44, 43, 41]) that the spine approach really began to crystalize. These papers laid out a formal basis for spines using a series of new measures on two underlying spaces of sample trees with and without distinguished lines of descent (spines). Of particular interest is the paper by Lyons [43] which gave a spine-based proof of the $\mathcal{L}^1$-convergence of the well-known martingale for the Galton-Watson process. Here we first saw the *spine decomposition* of the martingale as the key to using the intuition provided by Chauvin and Rouault's pathwise construction of the new measure – Lyons used this together with a previously known measure-theoretic result on Radon-Nikodym derivatives that allows us to deduce the behaviour of the change-of-measure martingale under the original measure by investigating its behaviour under the second measure. Similar ideas have recently been used by Kyprianou [40] to investigate the $\mathcal{L}^1$-convergence of the BBM martingale (1), by Biggins and Kyprianou [4] for multi-type branching processes in discrete time, by Hu and Shi [33] for the minimal position in a branching random walk, by Geiger [16, 17] for Galton-Watson processes, by Georgii and Baake [19] to study ancestral type behaviour in a continuous time branching Markov chain, as well as Olofsson [47] for general branching processes. Also see Athreya [2], Geiger [15, 18], Iksanov [34], Rouault and Liu [42] and Waymire and Williams [49], to name just a few other papers where spine and size-biasing techniques have already proved extremely useful in branching process situations. For applications of spines in branching in random media see, for example, the survey by Engländer [13].

In this article[1], we present a modified formalization of the spine approach that attempts to improve on the schemes originally laid out by Lyons *et al.* [44, 43, 41] and later for BBM by Kyprianou [40]. Although the set-up costs of our spine formalization are quite large, at least in terms of definitions and notation, the underlying ideas are all extremely simple and intuitive. One advantage of this approach is that it has facilitated the development of further spine techniques, for example, in Hardy & Harris [23, 22], Git *et al.* [20] and J.W.Harris & S.C.Harris [27] where a number of technical problems and difficult non-linear calculations are by-passed with spine calculations enabling their reduction to relatively straightforward classical one-particle situations; this article also serves as a foundation for these and other works.

The basic concept of our approach is quite straightforward: given the original branching process, we first create an extended probability measure by enriching the process through (carefully) choosing at random one of the particles to be the so-called *spine*. Now, on this enriched process, changes of measure can easily be applied that *only* affect the behaviour along the path of this single distinguished 'spine' particle; in our examples, we add a drift to the spine's motion, increase rate of fission along the path of the spine and size-bias the spine's offspring distribution. However, projecting this new enriched and changed measure back onto the original process filtration (that is, without any knowledge of the distinguished spine) brings the fundamental 'additive' martingales into play as a Radon-Nikodym derivative. The four probability measures, various martingales, extra filtrations and clear process constructions afforded by our setup, together with some other useful properties and tricks, such as the *spine decomposition*, provide a very elegant, intuitive and powerful set of techniques for analysing the process.

The reader who is familiar with the work of Lyons *et al.* [44] or Kyprianou [40] will notice significant similarities as well as differences in our approach. In the first instance our modifications correct our perceived weakness in the Lyons *et al.* scheme where one of the measures they defined had a time-dependent mass and could not be normalized to be a probability measure in a natural way, hence lacked a clear interpretation in terms of any direct process construction; an immediate consequence of this improvement is that here *all* measure changes are carried out by *martingales* and we regain a clear intuitive construction. Another difference is in our use of filtrations and sub-filtrations, where Lyons *et al.* instead used marginalizing. As we shall show, this brings substantial benefits since it allows us to relate the spine and the branching diffusion through the conditional expectation operation, and in this

[1] Based on the `arXiv` articles [24, 25]

way gives us a proper methodology for building *new* martingales for the branching diffusion based on known single particle martingales for the spine.

The conditional expectation approach also leads directly to simple proofs of some key results for branching diffusions. The first of these concerns the relation that becomes clear between the spine and the 'Gibbs-Boltzmann' weightings for the branching particles. Such weightings are well known in the theory of branching process, for example see Chauvin & Rouault [7], or Harris [30] which also studies the continuous-typed branching diffusion example introduced later. In our formulation these weightings can be interpreted as a conditional expectation of a spine event, and we can use them to immediately obtain a new interpretation of the additive operations previously seen only within the context of the Kesten-Stigum theorem and related problems. Our approach also leads to a substantially easier proof of a more general form of the *Many-to-One* theorem that is so often useful in branching processes applications; for example, in Champneys *et al.* [5] or Harris and Williams [28], special cases of this theorem were a key tool in their more classical approaches to branching diffusions.

As another application of spine techniques, we will analyze the $\mathcal{L}^p$-convergence properties (for $p \geq 1$) of some fundamental positive 'additive' martingales for three different models of branching diffusions.

Consider first the branching Brownian motion (BBM) with random family sizes. We recall that Kyprianou [40] used spine techniques to give necessary and sufficient conditions for $\mathcal{L}^1$-convergence of the Z_λ martingales:

Theorem 1.2 *Let $\tilde{\lambda} := -\sqrt{2rm}$ so that $c_\lambda := -E_\lambda/\lambda$ attains local maximum at $\tilde{\lambda}$. For each $x \in \mathbb{R}$, the limit $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ exists P^x-almost surely where:*

- *if $\lambda \leq \tilde{\lambda}$ then $Z_\lambda(\infty) = 0$ P^x-almost surely;*
- *if $\lambda \in (\tilde{\lambda}, 0]$ and $P(A \log^+ A) = \infty$ then $Z_\lambda(\infty) = 0$ P^x-almost surely;*
- *if $\lambda \in (\tilde{\lambda}, 0]$ and $P(A \log^+ A) < \infty$ then $Z_\lambda(t) \to Z_\lambda(\infty)$ almost surely and in $\mathcal{L}^1(P^x)$.*

(Note, without loss of generality (by symmetry) we will suppose $\lambda \leq 0$ throughout this article.)

In fact, in many cases where the martingale has a non-trivial limit, the convergence will also be much stronger than merely in $\mathcal{L}^1(P^x)$, as indicated by the following $\mathcal{L}^p$-convergence result:

Theorem 1.3 *For each $x \in \mathbb{R}$, and for each $p \in (1, 2]$:*

- *$Z_\lambda(t) \to Z_\lambda(\infty)$ almost surely and in $\mathcal{L}^p(P^x)$ if $p\lambda^2 < 2mr$ and $P(A^p) < \infty$*
- *Z_λ is unbounded in $\mathcal{L}^p(P^x)$, that is $\lim_{t\to\infty} P^x(Z_\lambda(t)^p) = \infty$, if $p\lambda^2 > 2mr$ or $P(A^p) = \infty$.*

We shall give a spine-based proof of this $\mathcal{L}^p$-convergence theorem, but also see Neveu [46] for sufficient conditions in the special case of binary branching at unit rate using more classical techniques. Also see Harris [29] for further discussion of martingale convergence in BBM and applications. Iksanov [34] also uses similar spine techniques in the study of the branching random walk.

For our second model, we look at a finite-type BBM model where the type of each particle controls the rate of fission, the offspring distribution and the spatial diffusion. First, we will extend Kyprianou's [40] approach to give the analogous $\mathcal{L}^1$-convergence result for this multi-type BBM model. We will also briefly discuss the rate of convergence of the martingales to zero and the speed of the spatially left-most particle within the process. Next, we give a new result on $\mathcal{L}^p$-convergence criteria, extending our earlier spine based proof developed for the single-type BBM case.

The third model we consider has a *continuous* type-space where the type of each particle moves independently as an Ornstein-Uhlenbeck process on $\mathbb{R}$. This branching diffusion was first introduced in Harris and Williams [28] and has also been investigated in Harris [30], Git *et al.* [20] and Kyprianou and Engländer [12].

Proofs for each of these models run along similar lines and the techniques are quite general, and it is a powerful feature of the spine approach that this is possible. For example, they have since been extended to more general branching diffusions in Engländer *et al.* [14] and to fragmentation processes in Harris *et al.* [31]. More classical techniques based on the expectation semigroup are simply not able to generalize easily, since they often require either some *a priori* bounds on the semigroup or involve difficult estimates – for example, in Harris and Williams [28] their important bound of a non-linear term is made possible only by the existence of a good $\mathcal{L}^2$ theory for their operator, and this is not generally available.

Of course, to prove martingale convergence in $\mathcal{L}^p$ for some $p > 1$ we use Doob's theorem, and therefore need only show that the martingale is *bounded* in $\mathcal{L}^p$. The *spine decomposition* is an excellent tool here for showing boundedness of the martingale since it reduces difficult calculations over the whole collection of branching particles to just the single spine process. We find the same conditions are also *necessary* for $\mathcal{L}^p$-boundedness of the martingale when $p > 1$ by just considering the contributions along the spine at times of fission and observing when these are unbounded. Otherwise, to determine whether the martingale is merely $\mathcal{L}^1$-convergent or has an almost-surely zero limit, we determine whether the martingale is almost-surely bounded or not under its own change of measure – this was Kyprianou's [40] approach and relies on a measure-theoretic result that has become standard in the spine methodology since the important work of Lyons *et al.* [44, 43, 41].

There are a number of reasons why we may be interested in knowing about the $\mathcal{L}^p$ convergence of a martingale: in Neveu's original article [46] it was a means to proving $\mathcal{L}^1$-convergence of martingales which can then be used to represent (non-trivial) travelling-wave solutions to the FKPP reaction-diffusion equation as well as in understanding the growth and spread of the BBM, whilst Git *et al.* [20] and Asmussen and Hering [1] have used it to deduce the almost-sure rate of convergence of the martingale to its limit. Of equal importance are the *techniques* that we use here. The convergence of other additive martingales can be determined with similar techniques, for example, see an application to a BBM with inhomogeneous breeding potential in J.W.Harris and S.C.Harris [27]. Similar ideas have also been used in proving a lower bound for a number of problems in the large-deviations theory of branching diffusions – we have used the spine decomposition with Doob's submartingale inequality to get an upper-bound for the growth of the martingale under the new measure which then leads to a lower-bound on the probability that one of the diffusing particles follows an unexpected path. See Hardy and Harris [23] for a spine-based proof of a path large deviation result for branching Brownian motion, and see Hardy and Harris [22] for a proof of a lower bound in the model that we consider in Section 11.

The layout of this paper is as follows. In Section 2, we will introduce the branching models, describing a binary branching multi-type BBM that we will frequently use as an example, before describing a more general branching Markov process model with random family sizes. In Section 3, we introduce the *spine* of the branching process as a distinguished infinite line of descent starting at the initial ancestor, we describe the underlying space for the branching Markov process with spine and we also introduce various fundamental filtrations. In Section 4, we define some fundamental probability spaces, including a probability measure for the branching process with a randomly chosen spine. In Section 5, various martingales are introduced and discussed. In particular, we see how to use filtrations and conditional expectation to build 'additive' martingales for the branching process out of the product of three simpler 'one-particle' martingales that only depend on the behaviour along the path of the spine; used as changes of measure, one martingale will increase the fission rate along the path of the spine, another will size-bias the offspring distribution along the spine, whilst the other one will change the motion of the spine. Section 6 discusses changes of measure with these martingales and gives very important and useful intuitive constructions for the branching process with spine under both the original measure $\tilde{P}$ and the changed measure $\tilde{\mathbb{Q}}$. Another extremely useful tool in the spine approach is the *spine decomposition* that we prove in Section 7; this gives an expression for the expectation of the 'additive' martingale under the new measure $\tilde{\mathbb{Q}}$ conditional on knowing the behaviour all along the path of the spine (including the spine's motion, the times of fission along the spine and number of offspring at each of the spine's fissions). In Section 8, we use the spine formulation to derive an interpretation for certain Gibbs-Boltzmann weights of $\tilde{\mathbb{Q}}$, discussing links with theorems of Kesten-Stigum and Watanabe, in addition to proving a 'Many-to-One' theorem. Finally, in sections 9, 10, and 11 we will prove the martingale convergence results for BBM, finite-type BBM and the continuous-type BBM models, respectively.

2 Branching Markov models

Before we present the underlying constructions for spines, it will be useful to give the reader a further idea of the branching-diffusion models that we have in mind for applications. We first briefly introduce a finite-type branching diffusion (which will often serve as a useful example), before presenting a more general model that shall be used as the basis of our spine constructions in the following sections.

2.1 A finite-type branching diffusion

Let θ be a strictly positive constant that can be considered as a temperature parameter. For some fixed $n \in \mathbb{N}$, define the finite type-space $I := \{1, \ldots, n\}$ and suppose that we are given two sets of positive constants $a(1), \ldots, a(n)$ and $R(1), \ldots, R(n)$.

A single particle motion. Consider the process $(\xi_t, \eta_t)_{t \geq 0}$ moving on $J := \mathbb{R} \times I$ as follows:

(i) The *type* location, η_t, of the particle moves as an irreducible, time-reversible Markov chain on the finite type-space I with Q-matrix θQ and invariant measure

$\pi = (\pi_1, \ldots, \pi_n)$;
(ii) the *spatial* location, ξ_t, moves as a driftless Brownian motion on $\mathbb{R}$ with diffusion coefficient $a(y) > 0$ whenever η_t is in state y, that is,

$$\mathrm{d}\xi_t = a(\eta_t)^{\frac{1}{2}}\,\mathrm{d}B_t, \qquad \text{where } B_t \text{ a Brownian motion.} \tag{3}$$

The formal generator of this process (ξ_t, η_t) is therefore:

$$\mathcal{H}F(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + \theta\sum_{j\in I} Q(y,j)F(x,j), \qquad (F : J \to \mathbb{R}). \tag{4}$$

A typed branching Brownian motion. Consider a branching diffusion where individual particles move independently according to the *single particle motion* as described above, and any particle currently of type y will undergo *binary* fission at rate $R(y)$ to be replaced by two particles at the same spatial and type positions as the parent. These offspring particles then move off independently, repeating stochastically the parent's behaviour, and so on.

Let the configuration of the whole branching diffusion at time t be given by the J-valued point process $\mathbb{X}_t = \{(X_u(t), Y_u(t)) : u \in N_t\}$, where N_t is the set of individuals alive at time t. Suppose probabilities for this process are given by $\{P^{x,y} : (x,y) \in J\}$ defined on the natural filtration, $(\mathcal{F}_t)_{t\geq 0}$, where $P^{x,y}$ is the law of the typed BBM process starting with one initial particle of type y at spatial position x.

This finite-type branching diffusion (with general offspring distribution) is investigated in Section 10 in this article, also see Hardy [21]. For now, we briefly introduce two fundamental *positive* martingales used to understand this model, the first based on the whole branching diffusion and the second based only on the single-particle model:

$$Z_\lambda(t) := \sum_{u\in N_t} v_\lambda(Y_u(t))e^{\lambda X_u(t) - E_\lambda t}, \tag{5}$$

$$\zeta_\lambda(t) := e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t}, \tag{6}$$

where v_λ and E_λ satisfy

$$\Big(\frac{1}{2}\lambda^2 A + \theta Q + R\Big)v_\lambda = E_\lambda v_\lambda,$$

where $A := \operatorname{diag}(a(y) : y \in I)$ and $R := \operatorname{diag}(R(y) : y \in I)$. That is, v_λ is the (Perron-Frobenius) eigenvector of the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$, with eigenvalue E_λ. These two martingales should be compared with the corresponding martingales (1) and $e^{\lambda B_t - \frac{1}{2}\lambda^2 t}$ for BBM and a single Brownian motion respectively.

2.2 A general branching Markov process

The spine constructions in our formulation can be applied to a much more general branching Markov model, and we shall base the presentation on the following model, where particles move independently in a general space J as a stochastic copy of some given Markov process Ξ_t, and at a location-dependent rate undergo fission to produce a location-dependent random number of offspring that each carry on this branching behaviour independently.

Definition 2.1 (A general branching Markov process) *We suppose that three initial elements are given to us:*

- *a Markov process Ξ_t in a measurable space $(J, \mathcal{B})$,*
- *a measurable function $R : J \to [0, \infty)$,*
- *for each $x \in J$ we are given a random variable $A(x)$ whose probability distribution on the numbers $\{0, 1, \dots\}$ is $P(A(x) = k) = p_k(x)$, with mean $m(x) := \sum_{k=0}^{\infty} k p_k(x) < \infty$.*

From these ingredients we can build a branching process in J according to the following recipe:

- *Each particle of the branching process will live, move and die in this space $(J, \mathcal{B})$, and if an individual u is alive at time t we refer to its location in J as $X_u(t)$. Therefore the time-t configuration of the branching process is a J-valued point process $\mathbb{X}_t := \{X_u(t) : u \in N_t\}$ where N_t denotes the collection of all particles alive at time t.*
- *For each individual u, the stochastic behaviour of its motion in J is an independent copy of the given process Ξ_t.*
- *The function $R : J \to [0, \infty)$ determines the rate at which each particle dies: given that u is alive at time t, its probability of dying in the interval $[t, t + \mathrm{d}t)$ is $R(X_u(t))\mathrm{d}t + o(\mathrm{d}t)$.*
- *If a particle u dies at location $x \in J$ it is replaced by $1 + A_u$ particles all positioned at x, where A_u is an independent copy of the random variable $A(x)$. All particles, once born, progress independently of each other.*

We suppose that the probabilities of this branching process are $\{P^x : x \in J\}$ where under P^x one initial ancestor starts out at x.

We shall first give a formal construction of the underlying probability space, made up of the sample trees of the branching process $\mathbb{X}_t$ in which the spines are the distinguished lines of descent. Once built, this space will be filtered in a natural way by the underlying family relationships of each sample tree, the diffusing branching particles and the diffusing spine, and then in section 4 we shall explain how we can define new probability measures $\tilde{P}^x$ that extend each P^x up to the finest filtration that contains all information about the spine and the branching particles. Much of the notation that we use for the underlying space of trees, the filtrations and the measures is closely related to that found in Kyprianou [40].

Although we do not strive to present our spine approach in the greatest possible generality, our model already covers many important situations whilst still being able to clearly demonstrate all the key spine ideas. In particular, in all our models, new offspring always inherit the position of their parent, although the same spine methods should also readily adapt to situations with random dispersal of offspring.

For greater clarity, we often use the finite-type branching diffusion of Section 2.1 to introduce the ideas before following up with the general formulation. For example, in this finite-type model we would take the process Ξ_t to be the single-particle process (ξ_t, η_t) which lives in the space $J := \mathbb{R} \times I$ and has generator $\mathcal{H}$ given by (4). The birth rate in this model at location $(x, y) \in J$ will be independent of x and given by the function $R(y)$ for all $y \in I$ and, since only binary branching occurs in this case, we also have $P(A(x, y) = 1) = 1$ for all $(x, y) \in J$.

3 The underlying space for spines

3.1 Marked Galton-Watson trees with spines

The set of Ulam-Harris labels is to be equated with the set Ω of finite sequences of strictly-positive integers:

$$\Omega := \{\emptyset\} \cup \bigcup_{n\in\mathbb{N}} (\mathbb{N})^n,$$

where we take $\mathbb{N} = \{1, 2, \ldots\}$. For two words $u, v \in \Omega$, uv denotes the concatenated word ($u\emptyset = \emptyset u = u$), and therefore Ω contains elements like '213' (or '$\emptyset$213'), which represents 'the 3rd child of the 1st child of the 2nd child of the initial ancestor $\emptyset$'. For two labels $v, u \in \Omega$ the notation $v < u$ means that v is an *ancestor* of u, and $|u|$ denotes the length of u. The set of all ancestors of u is equally given by

$$\{v : v < u\} = \{v : \exists w \in \Omega \text{ such that } vw = u\}.$$

Collections of labels, ie. subsets of Ω, will therefore be groups of individuals. In particular, a subset $\tau \subset \Omega$ will be called a *Galton-Watson tree* if:

1. $\emptyset \in \tau$,
2. if $u, v \in \Omega$, then $uv \in \tau$ implies $u \in \tau$,
3. for all $u \in \tau$, there exists $A_u \in 0, 1, 2, \ldots$ such that $uj \in \tau$ if and only if $1 \leq j \leq 1 + A_u$, (where $j \in \mathbb{N}$).

That is just to say that a Galton-Watson tree:

1. has a single initial ancestor $\emptyset$,
2. contains all ancestors of any of its individuals v,
3. has the $1 + A_u$ children of an individual u labelled in a consecutive way,

and is therefore just what we imagine by the picture of a family tree descending from a single ancestor. Note that the '$1 \leq j \leq 1 + A_u$' condition in 3 means that each individual has *at least* one child, so that in our model we are insisting that Galton-Watson trees *never die out*.

The set of all Galton-Watson trees will be called $\mathbb{T}$. Typically we use the name τ for a particular tree, and whenever possible we will use the letters u or v or w to refer to the labels in τ, which we may also refer to as *nodes of* τ or *individuals in* τ or just as *particles*.

Each individual should have a *location* in J at each moment of its *lifetime*. Since a Galton-Watson tree $\tau \in \mathbb{T}$ in itself can express only the *family* structure of the individuals in our branching random walk, in order to give them these extra features we suppose that each individual $u \in \tau$ has a mark (X_u, σ_u) associated with it which we read as:

- $\sigma_u \in \mathbb{R}^+$ is the *lifetime* of u, which determines the *fission time* of particle u as $S_u := \sum_{v\leq u} \sigma_v$ (with $S_\emptyset := \sigma_\emptyset$). The times S_u may also be referred to as the *death* times;
- $X_u : [S_u - \sigma_u, S_u) \to J$ gives the *location* of u at time $t \in [S_u - \sigma_u, S_u)$.

To avoid ambiguity, it is always necessary to decide whether a particle is in existence or not at its death time.

Remark 3.1 *Our convention throughout will be that a particle u dies 'just before' its death time S_u (which explains why we have defined $X_u : [S_u - \sigma_u, S_u) \to \cdot$ for example). Thus at the time S_u the particle u has* disappeared*, replaced by its $1 + A_u$ children which are all alive and ready to go.*

We denote a single marked tree by (τ, X, σ) or (τ, M) for shorthand, and the set of all marked Galton-Watson trees by $\mathcal{T}$:

- $\mathcal{T} := \Big\{(\tau, X, \sigma) : \tau \in \mathbb{T}$ and for each $u \in \tau, \sigma_u \in \mathbb{R}^+, X_u : [S_u - \sigma_u, S_u) \to J\Big\}$.
- For each $(\tau, X, \sigma) \in \mathcal{T}$, the set of particles that are alive at time t is defined as $N_t := \{u \in \tau : S_u - \sigma_u \le t < S_u\}$.

Where we want to highlight the fact that these values depend on the underlying marked tree we write e.g. $N_t((\tau, X, \sigma))$ or $S_u((\tau, M))$.

Any particle $u \in \tau$ that comes into existence creates a *subtree* made up from the collection of particles (and all their marks) that have u as an ancestor – and u is the original ancestor of this subtree.

- $(\tau, X, \sigma)^u_j$, or $(\tau, M)^u_j$ for shorthand, is defined as the *subtree* growing from individual u's jth child uj, where $1 \le j \le 1 + A_u$.

This subtree is a marked tree itself, but when considered as a part of the original tree we have to remember that it comes into existence at the space-time location $(X_u(S_u - \sigma_u), S_u - \sigma_u)$ – which is just the space-time location of the death of particle u (and therefore the space-time location of the birth of its child uj).

Before moving on there is a further useful extension of the notation: for any particle u we extend the definition of X_u from the time interval $[S_u - \sigma_u, S_u)$ to allow all earlier times $t \in [0, S_u)$:

Definition 3.2 *Each particle u is alive in the time interval $[S_u - \sigma_u, S_u)$, but we extend the concept of its path in J to all earlier times $t < S_u$:*

$$X_u(t) := \begin{cases} X_u(t) \text{ if } S_u - \sigma_u \le t < S_u \\ X_v(t) \text{ if } v < u \text{ and } S_v - \sigma_v \le t < S_v \end{cases}$$

Thus particle u inherits the path of its unique line of ancestors, and this simple extension will allow us to later write expressions like $\exp\{\int_0^t f(s)\,\mathrm{d}X_u(s)\}$ whenever $u \in N_t$, without worrying about the birth time of u.

For any given marked tree $(\tau, M) \in \mathcal{T}$ we can identify distinguished lines of descent from the initial ancestor: $\emptyset, u_1, u_2, u_3, \ldots \in \tau$, in which u_3 is a child of u_2, which itself is a child of u_1 which is a child of the original ancestor $\emptyset$. We'll call such a subset of τ a *spine*, and will refer to it as ξ:

- a *spine* ξ is a subset of nodes $\{\emptyset, u_1, u_2, u_3, \ldots\}$ in the tree τ that make up a unique line of descent. We use ξ_t to refer to the unique node in ξ that that is alive at time t.

In a more formal definition, which can for example be found in the paper by Rouault and Liu [42], a spine is thought of as a point on $\partial\tau$ the boundary of the tree – in fact the boundary is *defined* as the set of all infinite lines of descent. This explains the notation $\xi \in \partial\tau$ in the following definition: we augment the space $\mathcal{T}$ of marked trees to become

- $\tilde{\mathcal{T}} := \Big\{(\tau, M, \xi) : (\tau, M) \in \mathcal{T} \text{ and } \xi \in \partial\tau\Big\}$ is the set of *marked trees with distinguished spines.*

It is natural to speak of the *position of the spine at time t* which we think of as the position of the unique node that is in the spine and alive at time t:

- we define the time-t position of the spine as $\xi_t := X_u(t)$, where $u \in \xi \cap N_t$.

By using the notation ξ_t to refer to both the node in the tree and that node's spatial position we are introducing potential ambiguity. However, in practice the context will usually make clear which we intend, although if this is not the case we shall give the node a longer name:

- $\text{node}_t((\tau, M, \xi)) := u$ if $u \in \xi$ is the node in the spine alive at time t,

which may also be written as $\text{node}_t(\xi)$.

Finally, it will later be important to know how many fission times there have been in the spine, or what is the same, to know which generation of the family tree the node ξ_t is in (where the original ancestor $\emptyset$ is considered to be the 0th generation)

Definition 3.3 *We define the counting function*

$$n_t = \big|node_t(\xi)\big|,$$

which tells us which generation the spine node is in, or equivalently how many fission times there have been on the spine. For example, if $\xi_t = \big(\emptyset, u_1, u_2\big)$ *then both* $\emptyset$ *and* u_1 *have died and so* $n_t = 2$.

3.2 Filtrations

The reader who is already familiar with the Lyons *et al.* [41, 43, 44] papers will recall that they used two separate underlying spaces of marked trees *with* and *without* the spines, then marginalized out the spine when wanting to deal only with the branching particles as a whole. Instead, we are going to use the single underlying space $\tilde{\mathcal{T}}$, but define *four* filtrations of it that will encapsulate different knowledge.

Filtration $(\mathcal{F}_t)_{t\geq 0}$

We define a filtration of $\tilde{\mathcal{T}}$ made up of the σ-algebras:

$$\mathcal{F}_t := \sigma\Big((u, X_u, \sigma_u) : S_u \leq t \ ; \ (u, X_u(s) : s \in [S_u - \sigma_u, t]) : t \in [S_u - \sigma_u, S_u)\Big).$$

Then, $\mathcal{F}_t$ *knows everything that has happened to all the branching particles up to the time t, but does not know which one is the spine.* Each of these σ-algebras will be a subset of the limit defined as

$$\mathcal{F}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{F}_t\Big).$$

Filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$

In order to know about the spine, we make this filtration finer, defining $\tilde{\mathcal{F}}_t$ by adding into $\mathcal{F}_t$ the knowledge of which node is the spine at time t:

$$\tilde{\mathcal{F}}_t := \sigma\big(\mathcal{F}_t, \mathrm{node}_t(\xi)\big), \qquad \tilde{\mathcal{F}}_\infty := \sigma\Big(\bigcup_{t\geq 0} \tilde{\mathcal{F}}_t\Big).$$

Consequently, $\tilde{\mathcal{F}}_t$ *knows* everything *about the branching process and* everything *about the spine up to time* t, including which nodes make up the spine, when they were born, when they died (ie. the fission times S_u), and their family sizes.

Filtration $(\mathcal{G}_t)_{t\geq 0}$

We define a filtration of $\tilde{\mathcal{T}}$, $\{\mathcal{G}_t\}_{t\geq 0}$, which is generated by *only* the spatial motion of the spine by:

$$\mathcal{G}_t := \sigma\big(\xi_s : 0 \leq s \leq t\big), \qquad \mathcal{G}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{G}_t\Big),$$

Then, $\mathcal{G}_t$ *knows only about the spine's motion in* J *up to time* t, but does not actually know which line of descent in the family tree makes up the spine or anything about births along the spine.

Filtration $(\tilde{\mathcal{G}}_t)_{t\geq 0}$

We augment $\mathcal{G}_t$ by adding in information on the nodes that make up the spine (as we did from $\mathcal{F}_t$ to $\tilde{\mathcal{F}}_t$), as well as the knowledge of when the fission times occurred on the spine and how big the families were that were produced:

$$\tilde{\mathcal{G}}_t := \sigma\big(\mathcal{G}_t, (\mathrm{node}_s(\xi) : s \leq t), (A_u : u < \mathrm{node}_t(\xi))\big), \qquad \tilde{\mathcal{G}}_\infty := \sigma\Big(\bigcup_{t\geq 0} \tilde{\mathcal{G}}_t\Big).$$

Then, $\tilde{\mathcal{G}}_t$ *knows about* everything *along the spine up until time* t.

We note the obvious relationships between these filtrations of $\tilde{\mathcal{T}}$ that $\mathcal{F}_t \subset \tilde{\mathcal{F}}_t$ and $\mathcal{G}_t \subset \tilde{\mathcal{G}}_t \subset \tilde{\mathcal{F}}_t$. Trivially, we also note that $\mathcal{G}_t \nsubseteq \mathcal{F}_t$, since the filtration $\mathcal{F}_t$ does not know *which* line of descent makes up the spine.

4 Probability measures

Having now carefully defined the underlying space for our probabilities, we remind ourselves of the probability measures:

Definition 4.1 *For each $x \in J$, let P^x be the measure on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ such that the filtered probability space $(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}, P^\cdot)$ is the canonical model for $\mathbb{X}_t$, the branching Markov process described in Definition 2.1.*

For details of how the measures P^x are formally constructed on the underlying space of trees, we refer the reader to the work of Neveu [45] and Chauvin [8, 6]. Note, we could equally think of P^x as a measure on $(\mathcal{T}, \mathcal{F}_\infty)$, but it is convenient to use the enlarged sample space $\tilde{\mathcal{T}}$ for all our measure spaces, varying only the filtrations.

Our spine approach relies first on building a measure $\tilde{P}^x$ under which the spine is a single genealogical line of descent chosen from the underlying tree. If we are given a sample tree (τ, M) for the branching process, it is easy to verify that, if at each fission we make a uniform choice amongst the offspring to decide which line of descent continues the spine ξ, when $u \in \tau$ we have

$$\mathrm{Prob}(u \in \xi) = \prod_{v<u} \frac{1}{1+A_v}. \tag{7}$$

In the binary-branching case, for example, $\mathrm{Prob}(A_v = 1) = 1$ and then $\mathrm{Prob}(u \in \xi) = 2^{-|u|}$. This simple observation is the key to our method for extending the measures, and for this we make use of the following representation found in Lyons [43].

Theorem 4.2 *If $f \in m\tilde{\mathcal{F}}_t$, that is f is an $\tilde{\mathcal{F}}_t$-measurable function, then we can write:*

$$f = \sum_{u \in N_t} f_u \mathbf{1}_{(\xi_t = u)} \tag{8}$$

where $f_u \in m\mathcal{F}_t$.

As a simple example of this, in the case of the finite-typed branching diffusion of Section 2.1, such a representation would be:

$$e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t)\, e^{\lambda \xi_t - E_\lambda t} = \sum_{u \in N_t} e^{\int_0^t R(Y_u(s))\,\mathrm{d}s} v_\lambda(Y_u(t))\, e^{\lambda X_u(t) - E_\lambda t}\, \mathbf{1}_{(\xi_t = u)}. \tag{9}$$

Definition 4.3 *Given the measure P^x on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ we extend it to the probability measure $\tilde{P}^x$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by defining*

$$\int_{\tilde{\mathcal{T}}} f \, \mathrm{d}\tilde{P}^x := \int_{\tilde{\mathcal{T}}} \sum_{u \in N_t} f_u \prod_{v<u} \frac{1}{1+A_v} \, \mathrm{d}P^x, \tag{10}$$

for each $f \in m\tilde{\mathcal{F}}_t$ with representation like (8).

The previous approach to spines, exemplified in Lyons [43], used the idea of *fibres* to get a measure analogous to our $\tilde{P}$ that could measure the spine. However, a perceived weakness in this approach was that the corresponding measure had time-dependent total mass and could not be normalized to become a probability measure with an intuitive construction, unlike our $\tilde{P}$. Our idea of using the down-weighting term of (7) in the definition of $\tilde{P}$ is crucial in ensuring that we get a natural *probability* measure (look ahead to Lemma 4.9), and leads to the very useful situation in which *all* measure changes in our formulation are carried out by *martingales*.

Theorem 4.4 *This measure $\tilde{P}^x$ is an extension of P^x in that $P = \tilde{P}|_{\mathcal{F}_\infty}$.*

Proof: If $f \in m\mathcal{F}_t$ then the representation (8) is trivial and therefore by definition

$$\int_{\tilde{\mathcal{T}}} f \, \mathrm{d}\tilde{P} = \int_{\tilde{\mathcal{T}}} f \times \Big(\sum_{u \in N_t} \prod_{v<u} \frac{1}{1+A_v} \Big) \, \mathrm{d}P.$$

However, it can be seen that $\sum_{u \in N_t} \prod_{v<u} \frac{1}{1+A_v} = 1$ by retracing the sum back through the lines of ancestors to the original ancestor $\emptyset$, factoring out the product terms as each generation is passed. Thus

$$\int_{\tilde{\mathcal{T}}} f \, \mathrm{d}\tilde{P} = \int_{\tilde{\mathcal{T}}} f \, \mathrm{d}P.$$

□

Definition 4.5 *The filtered probability space $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty, (\tilde{\mathcal{F}}_t)_{t\geq 0}, \tilde{P})$ with $(\mathbb{X}_t, \xi_t)$ will be referred to as the* ***canonical model with spines****.*

In the single-particle model of section 2.1 we assumed the existence of a separate measure $\mathbb{P}$ and a process (ξ_t, η_t) that behaved stochastically like a 'typical' particle in the typed branching diffusion $\mathbb{X}_t$. In our formalization the *spine* is exactly the single-particle model:

Definition 4.6 *We define the measure $\mathbb{P}$ on $\big(\tilde{\mathcal{T}}, \mathcal{G}_\infty\big)$ as the restriction of $\tilde{P}$:*

$$\mathbb{P}|_{\mathcal{G}_t} := \tilde{P}|_{\mathcal{G}_t}.$$

Under the measure $\mathbb{P}$ the spine process ξ_t has exactly the same law as Ξ_t.

Definition 4.7 *The filtered probability space $(\tilde{\mathcal{T}}, \mathcal{G}_\infty, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P})$ together with the spine process ξ_t will be referred to as the* ***single-particle model****.*

4.1 An intuitive construction of $\tilde{P}$

As the name suggests, we should be able to think of the spine as the backbone of the branching process. This is made precise by the following decomposition:

Theorem 4.8 *The measure $\tilde{P}$ on $\tilde{\mathcal{F}}_t$ can be decomposed as:*

$$\mathrm{d}\tilde{P}(\tau, M, \xi) = \mathrm{d}\mathbb{P}(\xi)\mathrm{d}\mathbb{L}^{(R(\xi))}(n) \left(\prod_{v<\xi_t} \frac{1}{1+A_v} \right) \left(\prod_{v<\xi_t} p_{A_v}(\xi_{S_v}) \prod_{j=1}^{A_v} \mathrm{d}P\big((\tau, M)^v_j\big) \right) \tag{11}$$

where $\mathbb{L}^{(R(\xi))}$ is the law of the Poisson (Cox) process with rate $R(\xi_t)$ at time t, and we recall that $n = (n_t : t \geq 0)$ is the counting process of fission times along the spine.

We can summarise a clear intuitive picture of this decomposition in the following lemma:

Lemma 4.9 *The decomposition of measure $\tilde{P}$ at (11) enables the following construction:*

- *the spine's motion is determined by the single-particle measure $\mathbb{P}$;*
- *the spine undergoes fission at time t at rate $R(\xi_t)$;*
- *at the fission time of node v on the spine, the single spine particle is replaced by $1 + A_v$ children, with A_v being chosen independently and distributed according to the location-dependent random variable $A(\xi_{S_v})$ with probabilities $(p_k(\xi_{S_v}) : k = 0, 1, \ldots)$;*
- *the spine is chosen uniformly from the $1 + A_v$ children at the fission point v;*
- *each of the remaining A_v children gives rise to the independent subtrees $(\tau, M)^v_j$, for $1 \le j \le A_v$, which are not part of the spine and which are each determined by an independent copy of the original measure P shifted to their point and time of creation.*

5 Martingales

Starting with the single Markov process Ξ_t that lives in $(J, \mathcal{B})$ we have built $(\mathbb{X}_t, \xi_t)$, a branching Markov process with spines, in which the spine ξ_t behaves stochastically like the given Ξ_t. In this section we are going to show how *any* given martingale for the spine leads to a corresponding additive martingale for the whole branching model.

We have already seen an example of this for the finite-type model of section 2.1, when we introduced the two martingales:

$$Z_\lambda(t) := \sum_{u \in N_t} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t}, \qquad \zeta_\lambda(t) := e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t) e^{\lambda \xi_t - E_\lambda t}.$$

Just from their very form it has always been clear that they are closely related. What we shall later be demonstrating in full generality in this section is that the key to their relationship comes through generalising the following $\tilde{\mathcal{F}}_t$-measurable martingale for the multi-type BBM model:

Definition 5.1 *We define an $\tilde{\mathcal{F}}_t$-measurable martingale:*

$$\tilde{\zeta}_\lambda(t) := \prod_{u < \xi_t} (1 + A_u) \times v_\lambda(\eta_t) e^{\lambda \xi_t - E_\lambda t}. \tag{12}$$

An important result that we show in this article (Lemma 5.7) is that $Z_\lambda(t)$ and $\zeta_\lambda(t)$ are simply conditional expectations of this new martingale $\tilde{\zeta}_\lambda$. We emphasize that this relationship is only *possible* because of the construction of $\tilde{P}$ as a *probability* measure and using filtrations to capture the different knowledge generated by the spine and the branching particles. This idea of projection is also used in random fragmentation theory where it corresponds to the notion of tagged fragment, see Bertoin [3], for example.

Furthermore, in the general form that we present below it provides a consistent methodology for using well-known martingales for a single process ξ_t to get new additive martingales for the related branching process. In Hardy and Harris [23, 22] we use these powerful ideas to give substantially easier proofs of large-deviations problems in branching diffusions than have previously been possible.

Suppose that $\zeta(t)$ is a strictly positive $\big(\tilde{\mathcal{T}}, (\mathcal{G}_t)_{t\geq 0}, \tilde{P}\big)$-martingale, which is to say that it is a $\mathcal{G}_t$-measurable function that is a martingale with respect to the measure $\tilde{P}$. For example, in the case of our finite-type branching diffusion this could be the martingale $\zeta_\lambda(t)$ which is $\mathcal{G}_t$-measurable since it refers only to the spine process (ξ_t, η_t).

Definition 5.2 *We shall call $\zeta(t)$ a **single-particle martingale**, since it is $\mathcal{G}_t$-measurable and thus depends only to the spine ξ.*

Any such single-particle martingale can be used to define an additive martingale for the whole branching process via the representation (8):

Definition 5.3 *Suppose that we can represent the martingale $\zeta(t)$ as*

$$\zeta(t) = \sum_{u\in N_t} \zeta_u(t)\mathbf{1}_{(\xi_t=u)}, \tag{13}$$

for $\zeta_u(t) \in m\mathcal{F}_t$, as at (8). We can then define an $\mathcal{F}_t$-measurable process $Z(t)$ as

$$Z(t) := \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t),$$

*and refer to $Z(t)$ as the **branching-particle martingale**.*

The martingale property $Z(t)$ will be established in Lemma 5.7 after first building another martingale, $\tilde{\zeta}(t)$, from the single-particle martingale $\zeta(t)$. First, for clarity, we take a moment to discuss this definition of the additive martingale and the terms like $\zeta_u(t)$.

If we return to our familiar martingales (5) and (6), it is clear that

$$\zeta_\lambda(t) = e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t) e^{\lambda\xi_t - E_\lambda t} = \sum_{u\in N_t} e^{\int_0^t R(Y_u(s))\,\mathrm{d}s} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t}\,\mathbf{1}_{(\xi_t=u)}. \tag{14}$$

The 'ζ_u' terms of (13) could be here replaced with a more descriptive notation $\zeta_\lambda[(X_u, Y_u)](t)$, where

$$\zeta_u(t) = \zeta_\lambda[(X_u, Y_u)](t) := e^{\int_0^t R(Y_u(s))\,\mathrm{d}s} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t},$$

can be seen to essentially be a functional of the space-type path $(X_u(t), Y_u(t))$ of particle u. In this way the original single-particle martingale ζ_λ would be understood as a functional of the space-type path (ξ_t, η_t) of the spine itself and we could write

$$\zeta_\lambda(t) = \zeta_\lambda[(\xi, \eta)](t) = \sum_{u\in N_t} \zeta_\lambda[(X_u, Y_u)](t)\,\mathbf{1}_{(\xi_t=u)}.$$

This is the idea behind the representation (13), and in those typical cases where the single-particle martingale is essentially a functional of the paths of the spine ξ_t, as is the case for our $\zeta_\lambda(t)$, we should just think of ζ_u as being that same functional but evaluated over the path $X_u(t)$ of particle u rather than the spine ξ_t. The representation (13) can also be used as a more general way of treating other martingales that perhaps are not such a simple functional of the spine path.

Finally, from (14) it is clear that the additive martingale being defined by definition 5.3 is our familiar $Z_\lambda(t)$:

$$Z_\lambda(t) = \sum_{u \in N_t} e^{-\int_0^t R(Y_u(s))\,\mathrm{d}s} \zeta_\lambda[(X_u, Y_u)](t) = \sum_{u \in N_t} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t}.$$

Although definition 5.3 will work in general, in the main the spine approach is interested in martingales that can act as Radon-Nikodym derivatives between probability measures, and therefore we suppose from now on that $\zeta(t)$ is *strictly positive*, and therefore that the additive martingale $Z(t)$ is strictly positive.

The work of Lyons *et al.* [43, 41, 44], that of Chauvin and Rouault [9] and more recently of Kyprianou [40] suggests that when a change of measure is carried out with a branching-diffusion additive martingale like $Z(t)$ it is typical to expect three changes: the spine will gain a drift, its fission times will be increased and the distribution of its family sizes will be size-biased. In section 6.1 we shall confirm this, but we first take a separate look at the martingales that could perform these changes, and which we shall combine to obtain a martingale $\tilde{\zeta}(t)$ that will ultimately be used to change the measure $\tilde{P}$.

Theorem 5.4 *The expression*

$$\prod_{v < \xi_t} \big(1 + m(\xi_{S_v})\big) e^{-\int_0^t m(\xi_s) R(\xi_s)\,\mathrm{d}s}$$

is a $\tilde{P}$-martingale that will increase the rate at which fission times occur along the spine from $R(\xi_t)$ to $(1 + m(\xi_t))R(\xi_t)$:

$$\frac{\mathrm{d}\mathbb{L}_t^{((1+m(\xi))R(\xi))}}{\mathrm{d}\mathbb{L}_t^{(R(\xi))}} = \prod_{v < \xi_t} \big(1 + m(\xi_{S_v})\big) e^{-\int_0^t m(\xi_s) R(\xi_s)\,\mathrm{d}s}$$

where $\mathbb{L}^{(R(\xi))}$ is the law of the Poisson (Cox) process with rate $R(\xi_t)$ at time t.

Theorem 5.5 *The term*

$$\prod_{v < \xi_t} \frac{1 + A_v}{1 + m(\xi_{S_v})}$$

is a $\tilde{P}$-martingale that will change the measure by size-biasing the family sizes born from the spine:

$$\text{if } v < \xi_t, \text{ then} \qquad \mathit{Prob}(A_v = k) = \frac{(1+k)p_k(\xi_{S_v})}{1 + m(\xi_{S_v})}.$$

The proof of these two results is left as an easy exercise for the reader. The product of these two martingales with the single-particle martingale $\zeta(t)$ will simultaneously perform the three changes mentioned above:

Definition 5.6 *We define a $\tilde{\mathcal{F}}_t$-measurable martingale as*

$$\begin{aligned}\tilde{\zeta}(t) &:= \prod_{v<\xi_t}(1+A_v)e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \times \zeta(t) \\ &= \prod_{u<\xi_t}\frac{1+A_u}{1+m(\xi_{S_u})} \times \prod_{v<\xi_t}\big(1+m(\xi_{S_v})\big)e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \times \zeta(t). \qquad (15)\end{aligned}$$

Significantly, *only* the motion of the spine and the behaviour along the immediate path of the spine will be affected by any change of measure using this martingale. Also note, this martingale is the general form of $\tilde{\zeta}_\lambda(t)$ that we defined at (12) for our finite-type model.

The real importance of the size-biasing and fission-time-increase operations is that they introduce the correct terms into $\tilde{\zeta}(t)$ so that the following key relationships hold:

Lemma 5.7 *Both $Z(t)$ and $\zeta(t)$ are projections of $\tilde{\zeta}(t)$ onto their filtrations: for all $t \geq 0$,*

- $Z(t) = \tilde{P}\big(\tilde{\zeta}(t)\,|\,\mathcal{F}_t\big)$,
- $\zeta(t) = \tilde{P}\big(\tilde{\zeta}(t)\,|\,\mathcal{G}_t\big)$.

Proof: We use the representation (8) of $\tilde{\zeta}(t)$:

$$\tilde{\zeta}(t) = \sum_{u\in N_t}\prod_{v<u}(1+A_v)e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\mathbf{1}_{(\xi_t=u)}. \qquad (16)$$

Since $\tilde{P}\big(\mathbf{1}_{(\xi_t=u)}|\mathcal{F}_t\big) = \mathbf{1}_{(u\in N_t)} \times \prod_{v<u}(1+A_v)^{-1}$, it follows that

$$\begin{aligned}\tilde{P}\big(\tilde{\zeta}(t)|\mathcal{F}_t\big) &= \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t) \times \prod_{v<u}(1+A_v)\,\tilde{P}\big(\mathbf{1}_{(\xi_t=u)}|\mathcal{F}_t\big) \\ &= \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t) = Z(t).\end{aligned}$$

On the other hand, the martingale terms in (15) imply

$$\tilde{P}\big(\tilde{\zeta}(t)|\mathcal{G}_t\big) = \zeta(t) \times \tilde{P}\Big(\prod_{v<\xi_t}(1+A_v)e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\big|\mathcal{G}_t\Big) = \zeta(t).$$

□

6 Changing the measures

For the finite type model, the single-particle martingale $\zeta_\lambda(t)$ defined at (6) can be used to define a new measure for the single-particle model (as in [21]), via

$$\left.\frac{\mathrm{d}\mathbb{P}_\lambda}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = \frac{\zeta_\lambda(t)}{\zeta(0)}.$$

We have now seen the close relationships between the three martingales ζ_λ, Z_λ and $\tilde{\zeta}_\lambda$:

$$Z_\lambda(t) = \tilde{P}\big(\tilde{\zeta}_\lambda(t) \,|\, \mathcal{F}_t\big), \qquad \zeta_\lambda(t) = \tilde{P}\big(\tilde{\zeta}_\lambda(t) \,|\, \mathcal{G}_t\big),$$

and in this section we show in a more general form how these close relationships mean that a new measure $\tilde{\mathbb{Q}}_\lambda$ defined in terms of $\tilde{P}$ as

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = \frac{\tilde{\zeta}_\lambda(t)}{\tilde{\zeta}_\lambda(0)},$$

will induce measure changes on the sub-filtrations $\mathcal{G}_t$ and $\mathcal{F}_t$ of $\tilde{\mathcal{F}}_t$ whose Radon-Nikodym derivatives are given by $\zeta_\lambda(t)$ and $Z_\lambda(t)$ respectively. We will also give a useful intuitive construction of the measures $\tilde{P}$ and $\tilde{\mathbb{Q}}$.

Definition 6.1 *A measure $\tilde{\mathbb{Q}}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ is defined via its Radon-Nikodym derivative with respect to $\tilde{P}$:*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = \frac{\tilde{\zeta}(t)}{\tilde{\zeta}(0)}.$$

As we did for the measures P and $\mathbb{P}$ in Section 4, we can restrict $\tilde{\mathbb{Q}}$ to the sub-filtrations:

Definition 6.2 *We define the measure $\mathbb{Q}$ on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0})$ via*

$$\mathbb{Q} := \tilde{\mathbb{Q}}|_{\mathcal{F}_\infty}.$$

Definition 6.3 *We define the measure $\hat{\mathbb{P}}$ on $(\tilde{\mathcal{T}}, \mathcal{G}_\infty, (\mathcal{G}_t)_{t\geq 0})$ via*

$$\hat{\mathbb{P}} := \tilde{\mathbb{Q}}|_{\mathcal{G}_\infty}.$$

A consequence of our new formulation in terms of filtrations and the equalities of Lemma 5.7 is that the changes of measure are carried out by $Z(t)$ and $\zeta(t)$ on their subfiltrations:

Theorem 6.4

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\right|_{\mathcal{F}_t} = \frac{Z(t)}{Z(0)}, \qquad \text{and} \qquad \left.\frac{\mathrm{d}\hat{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = \frac{\zeta(t)}{\zeta(0)}.$$

Proof: These two results actually follow from a more general observation that if $\tilde{\mu}_1$ and $\tilde{\mu}_2$ are two measures defined on a measure space $(\Omega, \tilde{\mathcal{S}})$ with Radon-Nikodym derivative

$$\frac{\mathrm{d}\tilde{\mu}_2}{\mathrm{d}\tilde{\mu}_1} = f,$$

and if $\mathcal{S}$ is a sub-σ-algebra of $\tilde{\mathcal{S}}$, then the two measures $\mu_1 := \tilde{\mu}_1|_{\mathcal{S}}$ and $\mu_2 := \tilde{\mu}_2|_{\mathcal{S}}$ on $(\Omega, \mathcal{S})$ are related by the conditional expectation operation:

$$\frac{\mathrm{d}\mu_2}{\mathrm{d}\mu_1} = \tilde{\mu}_1\big(f|\mathcal{S}\big).$$

Applying this general result and using the relationships between the general martingales given in Lemma 5.7 concludes the proof. □

6.1 Understanding the measure $\tilde{\mathbb{Q}}$

This decomposition of $\tilde{P}_t$ given at (11) will allow us to interpret the measure $\tilde{\mathbb{Q}}$ if we appropriately factor the components of the change-of-measure martingale $\tilde{\zeta}(t)$ across this representation. On $\tilde{\mathcal{F}}_t$,

$$\begin{aligned}
\mathrm{d}\tilde{\mathbb{Q}} &= \tilde{\zeta}(t)\,\mathrm{d}\tilde{P}\\
&= \zeta(t) \times e^{-\int_0^t R(\xi_s)\mathrm{d}s} \prod_{u<\xi_t} \big(1 + m(\xi_{S_u})\big) \times \prod_{v<\xi_t} \frac{1 + A_v}{1 + m(\xi_{S_v})} \times \mathrm{d}\tilde{P}\\
&= \mathrm{d}\hat{\mathbb{P}}(\xi)\,\mathrm{d}\mathbb{L}^{((1+m(\xi))R(\xi))}(n)\\
&\qquad \times \prod_{u<\xi_t} \frac{1}{1 + A_u} \prod_{v<\xi_t} \frac{1 + A_v}{1 + m(\xi_{S_v})} p_{A_v}(\xi_{S_v}) \prod_{j=1}^{A_v} \mathrm{d}P\big((\tau, M)_j^v\big).
\end{aligned} \tag{17}$$

Just as we did for $\tilde{P}$, we can offer a clear interpretation of this decomposition:

Lemma 6.5 *Under the measure $\tilde{\mathbb{Q}}$,*

- *the spine process ξ_t moves as if under the* changed *measure $\hat{\mathbb{P}}$;*
- *the fission times along the spine occur at an* accelerated *rate $(1 + m(\xi_t))R(\xi_t)$;*
- *at the fission time of node v on the spine, the single spine particle is replaced by $1 + A_v$ children, with A_v being chosen as an independent copy of the random variable $\tilde{A}(y)$ which has the* size biased *offspring distribution $((1+k)p_k(y)/(1+m(y)) : k = 0, 1, \ldots)$, where $y = \xi_{S_v} \in J$ is the spine's location at the time of fission;*
- *the spine is chosen uniformly from the $1 + A_v$ particles at the fission point v;*
- *each of the remaining A_v children gives rise to the independent subtrees $(\tau, M)_j^v$, for $1 \le j \le A_v$, which are not part of the spine and evolve as independent processes determined by the measure P shifted to their point and time of creation.*

Such an interpretation of the measure $\tilde{\mathbb{Q}}$ was first given by Chauvin and Rouault [9] in the context of BBM, allowing them to come to the important conclusion that under the new measure $\mathbb{Q}$ the branching diffusion remains largely unaffected, except that the Brownian particles of a single (random) line of descent in the family tree are given a changed motion, with an accelerated birth rate – although they did not have random family sizes, so the size-biasing aspect was not seen. Size-biasing has been known for a long time in the study of branching populations, and in the context of spines, it was introduced in the Lyons *et al.* papers [43, 41, 44]. Kyprianou [40] presented the decomposition of equation (17) and the construction of $\mathbb{Q}$ at Lemma 6.5 for BBM with random family sizes, but did not follow our natural approach of starting with the *probability* measure $\tilde{P}$.

7 The spine decomposition

One of the most important results introduced in Lyons [43] was the so-called *spine decomposition*, which in the case of the additive martingale

$$Z_\lambda(t) = \sum_{u\in N_t} v_\lambda(Y_u(t))e^{\lambda X_u(t)-E_\lambda t},$$

from the finite-type branching diffusion would be:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{u<N_t} v_\lambda(\eta_{S_u})e^{\lambda\xi_{S_u}-E_\lambda S_u} + v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t}. \tag{18}$$

To prove this we start by decomposing the martingale as

$$Z_\lambda(t) = \sum_{u\in N_t, u\notin\xi} v_\lambda(Y_u(t))e^{\lambda X_u(t)-E_\lambda t} + v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t},$$

which is clearly true since one of the particles $u \in N_t$ must be in the line of descent that makes up the spine ξ. Recalling that the σ-algebra $\tilde{\mathcal{G}}_\infty$ contains all information about the line of nodes that makes up the spine, all about the spine diffusion (ξ_t, η_t) for all times t, and also contains all information regarding the fission times and number of offspring along the spine, it is useful to partition the particles $v \in \{u \in N_t, u \notin \xi\}$ into the distinct subtrees $(\tau, M)^u$ that were born at the fission times S_u from the particles that made up the spine before time t, or in other words those nodes in the $\{u < \xi_t\}$ of ancestors of the current spine node ξ_t. Thus:

$$Z_\lambda(t) = \sum_{u<\xi_t} e^{\lambda\xi_{S_u}-E_\lambda S_u}\Bigg\{\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\Bigg\} + v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t}.$$

If we now take the $\tilde{\mathbb{Q}}_\lambda$-conditional expectation of this, we find

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda&\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)\\ &= \sum_{u<\xi_t} e^{\lambda\xi_{S_u}-E_\lambda S_u}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\,\Big|\,\tilde{\mathcal{G}}_\infty\Big)\\ &\qquad + v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t}.\end{aligned}$$

We know from the decomposition (17) that the under the measure $\tilde{\mathbb{Q}}_\lambda$ the subtrees coming off the spine evolve as if under the measure P, and therefore

$$\tilde{\mathbb{Q}}_\lambda\Big(\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\big|\,\tilde{\mathcal{G}}_\infty\Big)$$
$$= \tilde{P}\Big(\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\big|\,\tilde{\mathcal{G}}_\infty\Big) = v_\lambda(\eta_{S_u}),$$

since the additive expression being evaluated on the subtree is just a shifted form of the martingale Z_λ itself.

This concludes the proof of (18), but before we go move on to give a similar proof for the general case, for easier reference through the cumbersome-looking general proof it is worth recalling that

$$\zeta_\lambda(t) = e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t) e^{\lambda\xi_t - E_\lambda t},$$

and therefore noting that (18) can alternatively be written as

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{u<N_t} e^{-\int_0^{S_u} R(\eta_s)\,\mathrm{d}s}\zeta_\lambda(S_u) + e^{-\int_0^t R(\eta_s)\,\mathrm{d}s}\zeta_\lambda(t).$$

Also, in the general model we are supposing that each particle u in the spine will give birth to a total of A_u subtrees that go off from the spine – the one remaining other offspring is used to continue the line of descent that makes up the spine. This explains the appearance of A_u in the general decomposition.

Theorem 7.1 (Spine decomposition) *We have the following **spine decomposition** for the additive branching-particle martingale:*

$$\tilde{\mathbb{Q}}\big(Z(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{u<\xi_t} A_u\, e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(S_u) + e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t).$$

Proof: In each sample tree one and only one of the particles alive at time t is the spine and therefore:

$$Z(t) = \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t),$$
$$= e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u\in N_t, u\neq\xi_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t).$$

The other individuals $\{u\in N_t, u\neq\xi_t\}$ can be partitioned into subtrees created from fissions along the spine. That is, each node u in the spine ξ_t (so $u<\xi_t$) has given birth at time S_u to one offspring node uj (for some $1\leq j\leq 1+A_u$) that was chosen to continue the spine whilst the other A_u individuals go off to make the subtrees $(\tau,M)^u_j$. Therefore,

$$Z(t) = e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u<\xi_t} e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s} \sum_{\substack{j=1,\ldots,1+A_u\\ uj\notin\xi}} Z_{uj}(S_u;t), \quad (19)$$

where for $t\geq S_u$,

$$Z_{uj}(S_u;t) := \sum_{v\in N_t, v\in(\tau,M)^u_j} e^{-\int_{S_u}^t m(X_v(s))R(X_v(s))\,\mathrm{d}s}\zeta_v(t),$$

is, conditional on $\tilde{\mathcal{G}}_\infty$, a $\tilde{P}$-martingale on the subtree $(\tau, M)^u_j$, and therefore

$$\tilde{P}\big(Z_{uj}(S_u;t)|\tilde{\mathcal{G}}_\infty\big) = \zeta(S_u).$$

Thus taking $\tilde{\mathbb{Q}}$-conditional expectations of (19) gives

$$\begin{aligned}\tilde{\mathbb{Q}}\big(Z(t)|\tilde{\mathcal{G}}_\infty\big) &= e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t)\\ &\quad + \sum_{u<\xi_t} e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}\tilde{P}\Big(\sum_{\substack{j=1,\ldots,1+A_u\\ uj\notin\xi}} Z_{uj}(S_u;t)\big|\,\tilde{\mathcal{G}}_\infty\Big)\\ &= e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u<\xi_t} e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}A_u\,\zeta(S_u),\end{aligned}$$

which completes the proof. □

This representation was first used in the Lyons *et al.* [43, 41, 44] papers and has become the standard way to investigate the behaviour of Z under the measure $\tilde{\mathbb{Q}}$. We also observe that the two measures $\tilde{P}$ and $\tilde{\mathbb{Q}}$ for the general model are equal when conditioned on $\tilde{\mathcal{G}}_\infty$ since this factors out their differences in the spine diffusion ξ_t, the family sizes born from the spine and the fission times on the spine. That is, $\tilde{P}\big(Z(t)|\tilde{\mathcal{G}}_\infty\big) = \tilde{\mathbb{Q}}\big(Z(t)|\tilde{\mathcal{G}}_\infty\big)$.

8 Spine results

Having covered the formal basis for our spine approach, we now present some results that follow from our spine formulation: the Gibbs-Boltzmann weights, conditional expectations, and a simpler proof of the improved Many-to-One theorem.

8.1 The *Gibbs-Boltzmann* weights of $\tilde{\mathbb{Q}}$

The Gibbs-Boltzmann weightings in branching processes are well-known, for example see Chauvin and Rouault [7] where they consider random measures on the boundary of the tree, and Harris [30] which gives convergence results for Gibbs-Boltzmann random measures. They have previously been considered via the individual terms of the additive martingale Z, but the following theorem gives a new interpretation of these weightings in terms of the spine. We recall that

$$Z(t) = \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t).$$

Theorem 8.1 *Let $u \in \Omega$ be a given and fixed label. Then*

$$\tilde{\mathbb{Q}}\big(\xi_t = u|\mathcal{F}_t\big) = \mathbf{1}_{(u\in N_t)}\frac{e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)}{Z(t)}.$$

Proof: Suppose $F \in \mathcal{F}_t$. We aim to show:

$$\int_F \mathbf{1}_{(\xi_t = u)} \, \mathrm{d}\tilde{\mathbb{Q}}(\tau, M, \xi) = \int_F \mathbf{1}_{(u \in N_t)} \frac{e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)}{Z(t)} \, \mathrm{d}\tilde{\mathbb{Q}}(\tau, M, \xi).$$

First of all we know that $\mathrm{d}\tilde{\mathbb{Q}}/\mathrm{d}\tilde{P} = \tilde{\zeta}(t)$ on $\mathcal{F}_t$ and therefore,

$$\mathrm{LHS} = \int_F \mathbf{1}_{(\xi_t = u)} \prod_{v < \xi_t} (1 + A_v) e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \zeta(t) \; \mathrm{d}\tilde{P}(\tau, M, \xi),$$

by definition of $\tilde{\zeta}(t)$ at (15). The definition 4.3 of the measure $\tilde{P}$ requires us to express the integrand with a representation like (8):

$$\mathbf{1}_{(\xi_t = u)} \prod_{v < \xi_t} (1 + A_v) e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \zeta(t)$$
$$= \mathbf{1}_{(\xi_t = u)} \mathbf{1}_{(u \in N_t)} \prod_{v < u} (1 + A_v) e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t),$$

and therefore

$$\begin{aligned}
\mathrm{LHS} &= \int_F \mathbf{1}_{(u \in N_t)} \prod_{v < u} (1 + A_v) e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \mathbf{1}_{(\xi_t = u)} \; \mathrm{d}\tilde{P}(\tau, M, \xi), \\
&= \int_F \mathbf{1}_{(u \in N_t)} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \; \mathrm{d}P(\tau, M, \xi),
\end{aligned}$$

by definition 4.3. Since $Z(t) > 0$ a.s., we know that on $\mathcal{F}_t$, $\mathrm{d}P/\mathrm{d}\mathbb{Q} = 1/Z(t)$, so

$$\mathrm{LHS} = \int_F \mathbf{1}_{(u \in N_t)} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \; \frac{1}{Z(t)} \mathrm{d}\mathbb{Q}(\tau, M, \xi),$$

and the proof is concluded. □

The above result combines with the representation (8) to show how we take conditional expectations under the measure $\tilde{\mathbb{Q}}$.

Theorem 8.2 *If $f(t) \in m\tilde{\mathcal{F}}_t$, and $f = \sum_{u \in N_t} f_u(t) \mathbf{1}_{(\xi_t = u)}$, with $f_u(t) \in m\mathcal{F}_t$ then*

$$\tilde{\mathbb{Q}}\big(f(t) | \mathcal{F}_t\big) = \sum_{u \in N_t} f_u(t) \frac{e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)}{Z(t)}. \tag{20}$$

Proof: It is clear that

$$\tilde{\mathbb{Q}}\big(f(t) | \mathcal{F}_t\big) = \sum_{u \in N_t} f_u(t) \tilde{\mathbb{Q}}\big(\xi_t = u | \mathcal{F}_t\big),$$

and the result follows from Theorem 8.1. □

A corollary to this useful result also appears to go a long way towards obtaining the Kesten-Stigum result in more general models:

Corollary 8.3 *If $g(\cdot)$ is a Borel function on J then*

$$\sum_{u \in N_t} g(X_u(t)) \, e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) = \tilde{\mathbb{Q}}\big(g(\xi_t) | \mathcal{F}_t\big) \times Z(t). \tag{21}$$

Proof: We can write $g(\xi_t) = \sum_{u\in N_t} g(X_u(t))\mathbf{1}(\xi_t = u)$, and now the result follows from the above corollary. □

The classical Kesten-Stigum theorems of [37, 36, 38] for multi-dimensional Galton-Watson processes give conditions under which an operation like the left-hand side of (21) converges as $t \to \infty$, and it is found that when it exists the limit is a multiple of the martingale limit $Z(\infty)$. Also see Lyons *et al.* [41] for a more recent proof of this based on other spine techniques. Our spine formulation apparently gives a previously unknown but simple meaning to this operation in terms of a conditional expectation and, as we hope to pursue in further work, in many cases we would intuitively expect that $\tilde{\mathbb{Q}}\big(g(\xi_t)|\mathcal{F}_t\big)/\tilde{\mathbb{Q}}\big(g(\xi_t)\big) \to 1$ a.s., leading to alternative spine proofs of both Kesten-Stigum like theorems and Watanabe's theorem in the case of BBM.

8.2 The *Full* Many-to-One Theorem

A very useful tool in the study of branching processes is the *Many-to-One* result that enables expectations of sums over particles in the branching process to be calculated in terms of an expectation of a single particle. In the context of the finite-type branching diffusion of section 2.1, the Many-to-One theorem would be stated as follows:

Theorem 8.4 *For any measurable function $f : J \to \mathbb{R}$ we have*

$$P^{x,y}\Big(\sum_{u\in N_t} f(X_u(t), Y_u(t))\Big) = \mathbb{P}^{x,y}\Big(e^{\int_0^t R(\eta_s)\,\mathrm{d}s} f(\xi_t, \eta_t)\Big).$$

Intuitively it is clear that the up-weighting term $e^{\int_0^t R(\eta_s)\,\mathrm{d}s}$ incorporates the notion of the population growing at an exponential rate, whilst the idea of $f(\xi_t, \eta_t)$ being the 'typical' behaviour of $f(X_u(t), Y_u(t))$ is also reasonable.

Existing results tend to apply only to functions of the above form that depend only on *the time-t location* of the spine and existing proofs do not lend themselves to covering functions that depend on the entire *path history* of the spine up to time t.

With the spine approach we have the benefit of being able to give a much less complicated proof of the stronger version that covers the most general path-dependent functions.

Theorem 8.5 (Many-to-One) *If $f(t) \in m\tilde{\mathcal{F}}_t$ has the representation*

$$f(t) = \sum_{u\in N_t} f_u(t)\mathbf{1}_{(\xi_t = u)},$$

where $f_u(t) \in m\mathcal{F}_t$, then

$$P\Big(\sum_{u\in N_t} f_u(t) e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\Big) = \tilde{P}\big(f(t)\,\tilde{\zeta}(t)\big) = \zeta(0)\,\tilde{\mathbb{Q}}\big(f(t)\big). \quad (22)$$

In particular, if $g(t) \in m\mathcal{G}_t$ with $g(t) = \sum_{u\in N_t} g_u(t)\mathbf{1}_{(\xi_t = u)}$ where $g_u(t) \in m\mathcal{F}_t$, then

$$P\Big(\sum_{u\in N_t} g_u(t)\Big) = \mathbb{P}\Big(e^{\int_0^t m(\xi_s)R(\xi_s)\mathrm{d}s} g(t)\Big) = \hat{\mathbb{P}}\Big(\frac{g(t)\,\zeta(0)}{e^{-\int_0^t m(\xi_s)R(\xi_s)\mathrm{d}s}\,\zeta(t)}\Big). \tag{23}$$

Proof: Let $f(t) \in m\tilde{\mathcal{F}}_t$ with the given representation. The tower property together with Theorem 8.2 gives

$$\begin{aligned}\tilde{\mathbb{Q}}\big(f(t)\big) &= \tilde{\mathbb{Q}}\Big(\tilde{\mathbb{Q}}\big(f(t)|\mathcal{F}_t\big)\Big) = \mathbb{Q}\Big(\tilde{\mathbb{Q}}\big(f(t)|\mathcal{F}_t\big)\Big)\\ &= \mathbb{Q}\Big(\frac{1}{Z(t)}\sum_{u\in N_t} f_u(t)e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\Big).\end{aligned}$$

From Theorem 6.4,

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\Big|_{\mathcal{F}_t} = \frac{Z(t)}{Z(0)},$$

and therefore we have

$$\tilde{\mathbb{Q}}\big(f(t)\big) = P\Big(Z(0)^{-1}\sum_{u\in N_t} f_u(t)e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\Big).$$

On the other hand,

$$\frac{\mathrm{d}\tilde{\mathbb{Q}}}{\mathrm{d}\tilde{P}}\Big|_{\tilde{\mathcal{F}}_t} = \frac{\tilde{\zeta}(t)}{\tilde{\zeta}(0)},$$

we have

$$\tilde{\mathbb{Q}}\big(f(t)\big) = \tilde{P}\big(f(t)\times\tilde{\zeta}(t)\,\tilde{\zeta}(0)^{-1}\big).$$

Trivially noting $Z(0) = \zeta(0) = \tilde{\zeta}(0)$ as there is only one initial ancestor, we can combine these expressions to obtain (22). For the second part, given $g(t) \in m\mathcal{G}_t$, we can define

$$f(t) := e^{\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} g(t)\times\zeta(t)^{-1},$$

which is clearly $\mathcal{G}_t$-measurable and satisfies $f(t) = \sum_{u\in N_t} f_u(t)\mathbf{1}_{(\xi_t=u)}$ with

$$f_u(t) = g_u(t)e^{\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)^{-1} \in m\mathcal{F}_t.$$

When we use this $f(t)$ in equation (22) and recall Lemma 5.7, that $\mathbb{P} := \tilde{P}|_{\mathcal{G}_\infty}$ from Definition 4.6 and that $\hat{\mathbb{P}} := \tilde{\mathbb{Q}}|_{\mathcal{G}_\infty}$ from Definition 6.3, we arrive at the particular case given at (23) in the theorem. □

In the further special case in which $g = g(\xi_t)$ for some Borel-measurable function $g(\cdot)$, the trivial representation

$$g(\xi_t) = \sum_{u\in N_t} g\big(X_u(t)\big)\mathbf{1}_{(\xi_t=u)}$$

leads immediately to the weaker version of the Many-to-One result that was utilised and proven, for example, in Harris and Williams [28] and Champneys *et al.* [5] using resolvents and the Feynman-Kac formula, expressed in terms of our more general branching Markov process $\mathbb{X}_t$:

Corollary 8.6 *If $g(\cdot) : J \to \mathbb{R}$ is $\mathcal{B}$-measurable then*

$$P\Big(\sum_{u\in N_t} g(X_u(t))\Big) = \mathbb{P}\Big(e^{\int_0^t R(\xi_s)\mathrm{d}s} g(\xi_t)\Big).$$

9 Branching Brownian motion

We now return to the original BBM model where particles move as standard Brownian motions, branching at rate r with offspring distribution A, as in Section 1. Under the measure $\tilde{P}^x$, the spine diffusion ξ_t is a Brownian motion that starts at x and we note that the martingale Z_λ can be obtained as in Sections 5 & 6 by starting with the spine $\tilde{P}^x$-martingale

$$\tilde{\zeta}_\lambda(t) := e^{-mrt}(1+m)^{n_t} \times \prod_{v<\xi_t} \left(\frac{1+A_v}{1+m}\right) \times e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}.$$

That is, we define the measure $\tilde{\mathbb{Q}}_\lambda$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}^x_\lambda}{\mathrm{d}\tilde{P}^x}\right|_{\tilde{\mathcal{F}}_t} := \frac{\tilde{\zeta}_\lambda(t)}{\tilde{\zeta}_\lambda(0)} = e^{\lambda(\xi_t - x) - E_\lambda t} \prod_{v<\xi_t} (1+A_v) \tag{24}$$

then, under $\tilde{\mathbb{Q}}^x_\lambda$, the process $\mathbb{X}_t$ can be constructed as follows:

- starting from x, the spine ξ_t diffuses according to a Brownian motion with drift λ on $\mathbb{R}$;
- at accelerated rate $(1+m)r$ the spine undergoes fission producing $1+\tilde{A}$ particles, where $\tilde{A}$ is independent of the spine's motion with size-biased distribution $\{(1+k)p_k/(1+m) : k \geq 0\}$;
- with equal probability, one of the spine's offspring particles is selected to continue the path of the spine, repeating stochastically the behaviour of its parent;
- the other particles initiate, from their birth position, independent copies of a $P^{\cdot}$ branching Brownian motion with branching rate r and family-size distribution given by A, that is, $\{p_k : k \geq 0\}$.

Further, ignoring information identifying the spine by setting $\mathbb{Q}^x_\lambda := \tilde{\mathbb{Q}}^x_\lambda|_{\mathcal{F}_\infty}$, we find

$$\left.\frac{\mathrm{d}\mathbb{Q}^x_\lambda}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \sum_{u\in N_t} e^{\lambda(X_u(t)-x) - E_\lambda t}. \tag{25}$$

Of course, this is all in full agreement with the equivalent definition of $\mathbb{Q}_\lambda$ initially introduced in Theorem 1.1 via its pathwise construction.

9.1 Proof of Theorem 1.3

Just before we proceed to the proof we recall the naturally occurring eigenvalue $E_\lambda := \frac{1}{2}\lambda^2 + mr$, noting that under the symmetry assumption that $\lambda \leq 0$ and for $p \in (1, 2]$:

$$pE_\lambda - E_{p\lambda} > 0 \quad \Leftrightarrow \quad c_\lambda > c_{p\lambda} \quad \Leftrightarrow \quad p\lambda^2 < 2mr$$

and that this always holds for some $p > 1$ whenever $\lambda \in (\tilde{\lambda}, 0]$, that is, when λ lies between the minimum of c_λ found at $\tilde{\lambda}$ and the origin.

Proof of part 1:

We are going to prove that for every $p \in (1, 2]$ the martingale Z_λ is $\mathcal{L}^p(P)$-convergent if $pE_\lambda - E_{p\lambda} > 0$. Furthermore, since $P^x(Z_\lambda(t)^p) = e^{p\lambda x} P^0(Z_\lambda(t)^p)$ we do not lose generality supposing that $x = 0$; from now on this is implicit if we drop the superscript by simply writing P.

From the change of measure at (25) it is clear that

$$P(Z_\lambda(t)^p) = P(Z_\lambda(t)^{p-1} Z_\lambda(t)) = \mathbb{Q}_\lambda(Z_\lambda(t)^q),$$

where $q := p - 1$. Our aim is to prove that $\mathbb{Q}_\lambda(Z_\lambda(t)^q)$ is bounded in t, since then $Z_\lambda^p(t)$ must be bounded in $\mathcal{L}^p(P)$ and Doob's theorem will then imply that Z_λ is convergent in $\mathcal{L}^p(P)$.

As we know from Theorem 7.1, the algebra $\tilde{\mathcal{G}}_\infty$ gives us the very important *spine-decomposition* of the martingale Z_λ:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{k=1}^{n_t} A_k e^{\lambda \xi_{S_k} - E_\lambda S_k} + e^{\lambda \xi_t - E_\lambda t}, \tag{26}$$

where A_k is the number of new particles produced from the fission at time S_k along the path of the spine, and the sum is taken to equal 0 if $n_t = 0$. The intuition is quite clear: since the particles that do not make up the spine grow to become independent copies of $\mathbb{X}_t$ distributed *as if under* P, the fact that Z_λ is a P-martingale on these subtrees implies that their contributions to the above decomposition are just equal to their *immediate* contribution on being born at time S_k at location ξ_{S_k}. Note, we emphasize that here we must use $\tilde{\mathbb{Q}}_\lambda$, since $\mathbb{Q}_\lambda$ cannot measure the algebra $\tilde{\mathcal{G}}_\infty \not\subseteq \mathcal{F}_\infty$.

We can now use the conditional form of Jensen's inequality followed by the spine decomposition of (26) coupled with the simple inequality,

Proposition 9.1 *If* $q \in (0, 1]$ *and* $u, v > 0$ *then* $(u + v)^q \le u^q + v^q$*,*

to obtain,

$$\begin{aligned} \tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)^q|\tilde{\mathcal{G}}_\infty\big) &\le \tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)^q & (27)\\ &\le \sum_{k=1}^{n_t} A_k^q e^{q\lambda \xi_{S_k} - qE_\lambda S_k} + e^{q\lambda \xi_t - qE_\lambda t}. & (28) \end{aligned}$$

With the tower property of conditional expectations and noting that $\mathbb{Q}_\lambda$ and $\tilde{\mathbb{Q}}_\lambda$ agree on $\mathcal{F}_t$,

$$\begin{aligned} \mathbb{Q}_\lambda(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)^q) &= \tilde{\mathbb{Q}}_\lambda\left(\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)^q|\tilde{\mathcal{G}}_\infty\big)\right) & (29)\\ &\le \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_k^q e^{q\lambda \xi_{S_k} - qE_\lambda S_k}\Big) + \tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_t - qE_\lambda t}\Big), & (30) \end{aligned}$$

and the proof of $\mathcal{L}^p(P)$-boundedness will be complete once we show this is bounded in t.

As written, (30) is made up of two terms, and since they play a central role from here on we name them explicitly: on the far right we have the **spine term** $\tilde{\mathbb{Q}}_\lambda\big(e^{q\lambda\xi_t - qE_\lambda t}\big)$, the other being the **sum term** $\tilde{\mathbb{Q}}_\lambda\big(\sum_{k=1}^{n_t} A_k^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\big)$.

The spine term: Changing from $\tilde{P}$ to $\tilde{\mathbb{Q}}_\lambda$ gives the spine a drift of λ, and therefore the change-of-measure for just the spine's motion (i.e. on the algebra $\mathcal{G}_t$) is carried out by the martingale $e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}$, so

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_t - qE_\lambda t}\Big) &= \tilde{P}\Big(e^{q\lambda\xi_t - qE_\lambda t} \times e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}\Big)\\ &= e^{\{\frac{1}{2}(p\lambda)^2 - \frac{1}{2}\lambda^2\}t - qE_\lambda t}\, \tilde{P}\Big(e^{p\lambda\xi_t - \frac{1}{2}(p\lambda)^2 t}\Big)\\ &= e^{-(pE_\lambda - E_{p\lambda})t}\, \tilde{\mathbb{Q}}_{p\lambda}(1) = e^{-(pE_\lambda - E_{p\lambda})t} \qquad (31)\end{aligned}$$

since the second-line term $e^{p\lambda\xi_t - \frac{1}{2}(p\lambda)^2 t}$ is also a $\tilde{P}$-martingale and $\frac{1}{2}(p\lambda)^2 - \frac{1}{2}\lambda^2 = E_{p\lambda} - E_\lambda$.

The sum term: Conditioning on the motion of the spine (without knowledge of the fission times or family sizes) and appealing to intuitive results from Poisson process theory (see [35] for example) yields

$$\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_k^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k} \,\big|\, \mathcal{G}_t\Big) = \int_0^t (1+m)r\, \tilde{\mathbb{Q}}_\lambda\big(\tilde{A}^q\big)\, e^{q\lambda\xi_s - qE_\lambda s}\, \mathrm{d}s \qquad (32)$$

Taking expectations of both sides of (32) and using Fubini's theorem then gives

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_k^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\Big) &= (1+m)r\, \tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) \int_0^t \tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_s - qE_\lambda s}\Big)\, \mathrm{d}s\\ &= (1+m)r\, \tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) \int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\, \mathrm{d}s, \qquad \text{using (31).}\end{aligned}$$

Thus we have found an explicit upper-bound (if $pE_\lambda \neq E_{p\lambda}$):

$$P^x(Z_\lambda(t)^p) \le e^{p\lambda x}\left(\frac{(1+m)r}{pE_\lambda - E_{p\lambda}}\Big[1 - e^{-(pE_\lambda - E_{p\lambda})t}\Big]\tilde{\mathbb{Q}}_\lambda\big(\tilde{A}^q\big) + e^{-(pE_\lambda - E_{p\lambda})t}\right). \qquad (33)$$

Finally, we also observe that

Lemma 9.2 *If $p \in (1,2]$ and $q := p-1$, $\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) < \infty$ if and only if $P(A^p) < \infty$*

since

$$\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) = \sum_{i=1}^{\infty} i^q \frac{i+1}{m+1} p_i = \frac{P(A^p) + P(A^q)}{m+1} \le \frac{2P(A^p)}{m+1}.$$

Hence, if we have $pE_\lambda - E_{p\lambda} > 0$ in addition to $P(A^p) < \infty$, this implies that $P^x(Z_\lambda(t)^p)$ will remain bounded as $t \to \infty$, which together with Doob's theorem will complete the proof of the first part of Theorem 1.3. □

Proof of Part 2:

We seek to show that Z_λ is unbounded in $\mathcal{L}^p(P^x)$ if either $pE_\lambda - E_{p\lambda} < 0$ or $P(A^p) = \infty$.

Note that *if* Z_λ is $\mathcal{L}^p(P^x)$ bounded then

$$P^x(Z_\lambda(\infty)^p) = \lim_{t\to\infty} P^x(Z_\lambda(t)^p) < \infty$$

hence, $\tilde{\mathbb{Q}}^x_\lambda(Z_\lambda(\infty)^q) < \infty$ and $Z_\lambda(\infty)^q$ is a uniformly integrable $\tilde{\mathbb{Q}}^x_\lambda$-submartingale. In particular, for any stopping time T, $\tilde{\mathbb{Q}}^x_\lambda\big(Z_\lambda(\infty)^q|\mathcal{F}_T\big) \geq Z_\lambda(T)^q$ hence $\tilde{\mathbb{Q}}^x_\lambda(Z_\lambda(\infty)^q) \geq \tilde{\mathbb{Q}}^x_\lambda(Z_\lambda(T)^q)$.

First, by considering only the contribution of the spine $Z_\lambda(t) \geq e^{\lambda\xi_t - E_\lambda t}$ for all $t \geq 0$ and recalling (31), we see that

$$\tilde{\mathbb{Q}}^x_\lambda(Z_\lambda(t)^q) \geq \tilde{\mathbb{Q}}^x_\lambda(e^{q\lambda\xi_t - qE_\lambda t}) = e^{q\lambda x - (pE_\lambda - E_{p\lambda})t}$$

and Z_λ is therefore unbounded in $\mathcal{L}^p(P^x)$ if $pE_\lambda - E_{p\lambda} < 0$.

Now, let T be any fission time along the path of the spine, then

$$Z_\lambda(T) \geq (1 + \tilde{A})e^{\lambda\xi_T - E_\lambda T}$$

where $\tilde{A}$ is the number of additional offspring produced at the time of fission. Then,

$$\tilde{\mathbb{Q}}^x_\lambda(Z_\lambda(T)^q) \geq \tilde{\mathbb{Q}}^x_\lambda\left((1 + \tilde{A})^q\right)\, e^{q\lambda x}\, \tilde{\mathbb{Q}}^x_{p\lambda}(e^{-(pE_\lambda - E_{p\lambda})T})$$

and so Z_λ is unbounded in $\mathcal{L}^p(P^x)$ if $\tilde{\mathbb{Q}}^x_\lambda\left((1 + \tilde{A})^q\right) = \infty$, which is true iff $P(\tilde{A}^p) = \infty$. $\square$

10 A typed branching diffusion

We move on to consider a general offspring distribution version of the typed branching diffusion introduced in Section 2.1. We will follow a similar notation and setup as before, but leave some details to the reader.

Recall the *single particle motion* $(X_t, Y_t)_{t\geq 0}$ from Section 2.1, where the *type* Y_t evolves as a Markov chain on $I := \{1, \ldots, n\}$ with Q-matrix θQ and the *spatial* location, X_t, moves as a driftless Brownian motion on $\mathbb{R}$ with diffusion coefficient $a(y) > 0$ whenever η_t is in state y.

Consider a *typed branching Brownian motion* where individual particles move independently according to the single particle motion as above, and any particle currently of type y will undergo fission at rate $R(y)$ to be replaced by a random number of offspring, $1 + A(y)$, where $A(y) \in \big\{0, 1, 2, \ldots\big\}$ is an independent RV with distribution

$$P\big(A(y) = i\big) = p_i(y), \qquad i \in \big\{0, 1, \ldots\big\},$$

and mean $M(y) := P\big(A(y)\big) < \infty$ for all $y \in I$. At birth, offspring inherit the parent's spatial and type positions and then move off independently, repeating stochastically the parent's behaviour, and so on. We gather together the mean number of offspring in matrix $M := \mathrm{diag}[M(1), \ldots, M(n)]$ and also recall that $R := \mathrm{diag}[R(1), \ldots, R(n)]$ and $A := \mathrm{diag}[a(1), \ldots, a(n)]$.

As usual, let the configuration of the whole branching diffusion at time t be given by the J-valued point process $\mathbb{X}_t = \{(X_u(t), Y_u(t)) : u \in N_t\}$, where N_t is the set of individuals alive at time t. Let the probabilities for this process be given by $\{P^{x,y} : (x, y) \in J\}$ defined on the natural filtration, $(\mathcal{F}_t)_{t\geq 0}$, where $P^{x,y}$ is the law of the typed BBM process starting with one initial particle of type y at spatial position x. Recall, under the extended measures $\{\tilde{P}^{x,y} : (x, y) \in J\}$ where we identify a distinguished infinite line of decent starting from the initial particle, this *spine* $(\xi_t, \eta_t)_{t\geq 0}$ will simply move like the *single particle motion* above.

It should be noted that the condition of time-reversibility on the Markov chain is not absolutely necessary, and is really just a simplifying assumption that gives us an easier $\mathcal{L}^2$ theory for the matrices and eigenvectors; our aim is really to show how the spine techniques work – lessening the geometric complexity of the model serves a good purpose.

Note, the special case of the 2-type BBM model was considered in Champneys *et al.* [5] by different means. Also, in our model, at the time of fission a type-y individual can produce only type-y offspring. This is not the same as the case in which a type-y individual may produce a random collection of particles of *different* types – as considered in T.E. Harris's classic text [32], for example. Other forms of typed branching processes have also been dealt with by spine techniques, for example, see Lyons *et al.* [41] or Athreya [2] for discrete-time models in which a particle's type does not change during its life but a type-w individual can give offspring of any type according to some distribution. See also the remarkable work of Georgii and Baake [19] that uses spine techniques to study ancestral type behaviour in a continuous time branching Markov chain where particles can give birth to across all types. In principle, our spine methods will be robust enough to extend to all these other type behaviours (with added spatial diffusion).

10.1 The martingale

Via the many-to-one Lemma 8.5, it is easy to see that for any $\lambda \in \mathbb{R}$, any function (vector) $v_\lambda : I \to \mathbb{R}$ and any number $E_\lambda \in \mathbb{R}$, the expression

$$Z_\lambda(t) := \sum_{u\in N(t)} v_\lambda(Y_u(t))\, e^{\lambda X_u(t) - E_\lambda t},$$

will be a martingale if and only if v_λ and E_λ satisfy:

$$\Big(\frac{1}{2}\lambda^2 A + \theta Q + MR\Big) v_\lambda = E_\lambda v_\lambda. \tag{34}$$

That is, v_λ must be an eigenvector of the matrix $\frac{1}{2}\lambda^2 A + \theta Q + MR$, with eigenvalue E_λ.

Definition 10.1 *For two vectors u, v on I, we define*

$$\langle u, v \rangle_\pi := \sum_{i=1}^{n} u_i v_i \pi_i,$$

which gives us a Hilbert space which we refer to as $\mathcal{L}^2(\pi)$. *We suppose that the eigenvector* v_λ *is normalized so that* $\|v_\lambda\|_\pi := \langle v_\lambda, v_\lambda \rangle_\pi = 1$.

The fact that the Markov chain is time-reversible implies that the matrix $\frac{1}{2}\lambda^2 A + \theta Q + MR$ is self-adjoint with respect to this inner product. This in itself is enough to guarantee the existence of eigenvectors in $\mathcal{L}^2(\pi)$, but the fact that we are dealing with a finite-state Markov chain means that we also have the *Perron-Frobenius* theory to hand, which allows us to suppose that v_λ is a *strictly positive* eigenvector whose eigenvalue E_λ is real and the farthest to the right of all the other eigenvalues – see Seneta [48] for details. This implies a useful representation for the eigenvalue:

Theorem 10.2

$$E_\lambda = \sup_{\|v\|_\pi = 1} \left\langle \left((\lambda^2/2)A + \theta Q + MR\right) v, v \right\rangle_\pi, \tag{35}$$

since it is the rightmost eigenvalue.

A proof can be found in Kreyzig [39]. From this it is not difficult to show that E_λ is a strictly-convex function of λ. Interestingly, it will be seen in our proofs that it is the geometry of the eigenvalue E_λ that determines the interval that gives rise to martingales $Z_\lambda(t)$ that are $\mathcal{L}^p$-convergent.

Corollary 10.3 *As a function of* λ, E_λ *is strictly-convex and infinitely differentiable with*

$$E'_\lambda = \lambda \langle A v_\lambda, v_\lambda \rangle_\pi. \tag{36}$$

If we define the speed function

$$c_\lambda := -E_\lambda/\lambda, \tag{37}$$

then on $(-\infty, 0)$ *the function* c_λ *has just one minimum at a single point* $\tilde{\lambda}(\theta)$, *either side of which* c_λ *is strictly increasing to* $+\infty$ *as either* $\lambda \downarrow -\infty$ *or* $\lambda \uparrow 0$. *In particular, for each* $\lambda \in (\tilde{\lambda}(\theta), 0]$ *there is some* $p > 1$ *such that* $c_\lambda > c_{p\lambda}$*; on the other hand, if* $\lambda < \tilde{\lambda}(\theta)$ *there is no such* $p > 1$.

We refer to the function c_λ as the speed function since it relates to the asymptotic speed of the travelling waves associated with the martingale $Z_\lambda(t)$; see Harris [29] or Champneys *et al.* [5] for details of the relationship between branching-diffusion martingales and travelling waves.

Since $Z_\lambda(t)$ is a strictly-positive martingale it is immediate that $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ exists and is finite almost-surely under $P^{x,y}$. As before, *by symmetry we shall assume that* $\lambda \leq 0$ *and, without loss of generality, we also suppose that* $P(A(y) = 0) = 1$ *whenever* $r(y) = 0$ *to simplify statements*. We shall prove necessary and sufficient conditions for $\mathcal{L}^1$-convergence of the Z_λ martingales:

Theorem 10.4 *For each* $x \in \mathbb{R}$, *the limit* $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ *exists* $P^{x,y}$*-a.s. where:*

- *if* $\lambda \le \tilde\lambda(\theta)$ *then* $Z_\lambda(\infty) = 0$ $P^{x,y}$*-almost surely;*
- *if* $\lambda \in (\tilde\lambda(\theta), 0]$ *and* $P(A(y)\log^+ A(y)) = \infty$ *for some* $y \in I$*, then* $Z_\lambda(\infty) = 0$ $P^{x,y}$*-a.s.;*
- *if* $\lambda \in (\tilde\lambda(\theta), 0]$ *and* $P(A(y)\log^+ A(y)) < \infty$ *for all* $y \in I$*, then* $Z_\lambda(t) \to Z_\lambda(\infty)$ *almost surely and in* $\mathcal{L}^1(P^{x,y})$*.*

Once again, in many cases where the martingale has a non-trivial limit, the convergence will be much stronger than merely in $\mathcal{L}^1(P^{x,y})$, as indicated by the following new $\mathcal{L}^p$-convergence result that we will prove by extending our earlier new spine approach:

Theorem 10.5 *For each* $x \in \mathbb{R}$*, and for each* $p \in (1, 2]$*:*

- $Z_\lambda(t) \to Z_\lambda(\infty)$ *a.s. and in* $\mathcal{L}^p(P^{x,y})$ *if* $pE_\lambda - E_{p\lambda} > 0$ *and* $P(A(y)^p) < \infty$ *for all* $y \in I$*.*
- Z_λ *is unbounded in* $\mathcal{L}^p(P^{x,y})$*, that is* $\lim_{t\to\infty} P^{x,y}(Z_\lambda(t)^p) = \infty$*, if either* $pE_\lambda - E_{p\lambda} < 0$ *or* $P(A(y)^p) = \infty$ *for some* $y \in I$*.*

Note, when $\lambda \le 0$*, the inequality* $pE_\lambda - E_{p\lambda} > 0$ *is equivalent to* $c_\lambda > c_{p\lambda}$ *and holds for some* $p \in (1, 2]$ *if and only if* $\lambda \in (\tilde\lambda(\theta), 0]$*.*

10.2 New measures for the typed BBM

As usual, we can define a measure $\tilde{\mathbb{Q}}_\lambda$ via a Radon-Nikodym derivative with respect to $\tilde P$ by combining three simpler changes of measures that only affect behaviour along the spine.

First, we observe that for $\lambda \in \mathbb{R}$,

$$v_\lambda(\eta_t)\, e^{\int_0^t MR(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}$$

is a $\tilde P$-martingale. This fact is easy to confirm with some classical 'one-particle' calculations, for example, using the Feynman-Kac formula, the generator (4) and noting the relation (34).

We can obtain the Z_λ martingale as in Sections 5 & 6 by using the $\tilde P$-martingale

$$\tilde\zeta_\lambda(t) := e^{-\int_0^t MR(\eta_s)\,\mathrm{d}s} \prod_{u<\xi_t} \big(1 + A_u\big) \times v_\lambda(\eta_t)\, e^{\int_0^t MR(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}.$$

That is, for each $\lambda \in \mathbb{R}$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde P^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{\tilde\zeta_\lambda(t)}{\tilde\zeta_\lambda(0)} = \frac{v_\lambda(\eta_t)}{v_\lambda(y)}\, e^{\lambda(\xi_t - x) - E_\lambda t} \prod_{v<\xi_t} (1 + A_v) \tag{38}$$

and then ignoring information about the spine by defining $\mathbb{Q}_\lambda^{x,y} := \tilde{\mathbb{Q}}_\lambda^{x,y}|_{\mathcal{F}_\infty}$, we find that

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = v_\lambda(y)^{-1} \sum_{u\in N(t)} v_\lambda(Y_u(t))\, e^{\lambda(X_u(t)-x) - E_\lambda t}. \tag{39}$$

We emphasise that, starting with the three simple 'spine' martingales, we have actually shown that Z_λ must, in fact, be a martingale. This route offers a simple way of getting general 'additive' martingales for the branching process.

10.3 The spine process (ξ_t, η_t) under $\tilde{\mathbb{Q}}_\lambda$

It remains to identify the behaviour of the spine under the change of measure. In the BBM model it was clear to see that the spine ξ_t received a drift under the measure $\tilde{\mathbb{Q}}_\lambda$, and something similar happens here:

Lemma 10.6 *Under $\tilde{\mathbb{Q}}_\lambda$ the spine process (ξ_t, η_t) has generator:*

$$\mathcal{H}_\lambda F(x,y) := \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + a(y)\lambda\frac{\partial F}{\partial x} + \sum_{j\in I}\theta Q_\lambda(y,j)F(x,j), \qquad (40)$$

where Q_λ is an honest Q-matrix:

$$\theta Q_\lambda(i,j) = \begin{cases} \theta Q(i,j)\frac{v_\lambda(j)}{v_\lambda(i)} & \text{if } i \neq j \\ \theta Q(i,i) + \frac{\lambda^2}{2}a(i) - E_\lambda + r(i) & \text{if } i = j \end{cases}$$

That is, under $\tilde{\mathbb{Q}}_\lambda$, ξ_t is a Brownian motion with instantaneous variance $a(\eta_t)$ and instantaneous drift $a(\eta_t)\lambda$, and η_t is a Markov chain on I with Q-matrix θQ_λ and invariant measure $\pi_\lambda = v_\lambda^2\pi$.

The form of this above generator $\mathcal{H}_\lambda$ can be obtained from the theory of Doob's h-transforms, due to the fact that on the algebra $\mathcal{G}_t$ the change of measure is given by:

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{G}_t} = \frac{1}{v_\lambda(y)e^{\lambda x}}\, v_\lambda(\eta_t)\, e^{\int_0^t MR(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}. \qquad (41)$$

The long-term behaviour under $\tilde{\mathbb{Q}}_\lambda$ of the spine diffusion ξ_t can now be retrieved from the generator (40) and the properties of E_λ stated in Lemma 10.3:

Corollary 10.7 *Almost surely under $\tilde{\mathbb{Q}}_\lambda^{x,y}$, the long-term drift of the spine is given explicitly as*

$$\lim_{t\to\infty} t^{-1}\xi_t = E'_\lambda$$

and hence

$$\xi_t + c_\lambda t \to \begin{cases} \infty & \text{if } \lambda \in (\tilde{\lambda}(\theta), 0] \\ -\infty & \text{if } \lambda < \tilde{\lambda}(\theta) \end{cases} \qquad (42)$$

whereas, if $\lambda = \tilde{\lambda}$ the process $\xi_t + c_\lambda t$ will be recurrent on $\mathbb{R}$ under $\tilde{\mathbb{Q}}_\lambda$.

Proof: From the generator stated at (40) we can write:

$$\xi_t = B\left(\int_0^t a(\eta_s)\mathrm{d}s\right) + \lambda\int_0^t a(\eta_s)\mathrm{d}s,$$

where $B(t)$ is a $\tilde{\mathbb{Q}}_\lambda$-Brownian motion. Then by the ergodic theorem and the fact that $\pi_\lambda = v_\lambda^2\pi$:

$$t^{-1}\xi_t \to \lambda\sum_{y\in I}a(y)\pi_\lambda(y) = \lambda\sum_{y\in I}a(y)v_\lambda^2(y)\pi(y) = \lambda\langle Av_\lambda, v_\lambda\rangle_\pi = E'_\lambda.$$

Direct calculation from (37) gives $E'_\lambda = -c_\lambda - \lambda c'_\lambda$, and therefore $t^{-1}(\xi_t + c_\lambda t) \to -\lambda c'_\lambda$, whence whether we are to the left or right of the local minimum of c_λ found at $\tilde{\lambda}$ determines the behaviour of $\xi_t + c_\lambda t$, as is required. Lastly, when $\lambda = \tilde{\lambda}$, with the laws of the iterated logarithm in mind, it is not difficult to see that both $B(\int_0^t a(\eta_s)\,ds)$ and $\int_0^t (\lambda a(\eta_s) + c_\lambda)\,ds$ will fluctuate about the origin, hence $\xi_t + c_\lambda t$ will be recurrent under $\tilde{\mathbb{Q}}_\lambda$. □

10.4 Construction of the process under $\tilde{\mathbb{Q}}_\lambda$

Drawing together the elements from this section, we now present the pathwise construction of the new measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$:

Theorem 10.8 *Under $\tilde{\mathbb{Q}}_\lambda^{x,y}$, the process $\mathbb{X}_t$ evolves as follows:*

- *starting from (x, y), the spine (ξ_t, η_t) evolves as a Markov process with generator $\mathcal{H}_\lambda$, that is, η_t evolves as Markov chain on I with Q-matrix θQ_λ and ξ_t moves as a Brownian motion on $\mathbb{R}$ with variance coefficient $a(\eta_t)$ and drift $a(\eta_t)\lambda$.*
- *whenever the type of the spine η is in state $y \in I$, the spine undergoes fission at an accelerated rate $(1 + m(y))r$, producing $1 + \tilde{A}(y)$ particles where $\tilde{A}(y)$ is independent of the spine's motion with size-biased distribution $\{(1+k)p_k(y)/(1+m(y)) : k \geq 0\}$;*
- *with equal probability, one of the spine's offspring particles is selected to continue the path of the spine, repeating stochastically the behaviour of its parent;*
- *the other particles initiate, from their birth position, independent copies of $P^{\cdot,\cdot}$ typed branching Brownian motions.*

10.5 Proof of Theorem 10.4

The following proof is an extension of that given for BBM by Kyprianou [40]. The second part of the following theorem is the key element in using the measure change (38) to determine properties of the martingale Z_λ:

Theorem 10.9 *Suppose that P and $\mathbb{Q}$ are two probability measures on a space $(\Omega, \mathcal{F}_\infty)$ with filtration $(\mathcal{F}_t)_{t\geq 0}$, such that for some positive martingale Z_t,*

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\right|_{\mathcal{F}_t} = Z_t.$$

The limit $Z_\infty := \limsup_{t\to\infty} Z_t$ therefore exists and is finite almost surely under P. Furthermore, for any $F \in \mathcal{F}_\infty$

$$\mathbb{Q}(F) = \int_F Z_\infty \,\mathrm{d}P + \mathbb{Q}\big(F \cap \{Z_\infty = \infty\}\big), \tag{43}$$

and consequently

$$(a) \quad P(Z_\infty = 0) = 1 \iff \mathbb{Q}(Z_\infty = \infty) = 1 \tag{44}$$

$$(b) \quad P(Z_\infty) = 1 \iff \mathbb{Q}(Z_\infty < \infty) = 1 \tag{45}$$

A proof of the decomposition (43) can be found in Durrett [11], at page 241.

Suppose that $\lambda \le \tilde{\lambda} < 0$. Ignoring all contributions except for the spine, it is immediate that

$$Z_\lambda(t) = \sum_{u \in N_t} v_\lambda(Y_u(t))\, e^{\lambda X_u(t) - E_\lambda t} \ge v_\lambda(\eta_t)\, e^{\lambda(\xi_t + c_\lambda t)}$$

where, from Corollary 10.7, under the measure $\tilde{\mathbb{Q}}_\lambda$ the spine satisfies $\liminf\{\xi_t + c_\lambda t\} = -\infty$ a.s. and $v_\lambda > 0$, hence $\limsup_{t\to\infty} Z_\lambda(t) = \infty$ almost surely under $\tilde{\mathbb{Q}}_\lambda$, yielding $P(Z_\lambda(\infty) = 0) = 1$.

Note that, for $y \in I$, $P(A(y)\log^+ A) < \infty \iff \sum_{k\ge1} P(\log^+ \tilde{A}(y) > ck) < \infty$ for any $c > 0$, where recall that $\tilde{A}(y)$ has the size-biased distribution $\{(i+1)p_k(y)/(1+m(y)) : k \ge 0\}$. Then for an IID sequence $\{\tilde{A}_n(y)\}$ of copies of $\tilde{A}(y)$, Borel-Cantelli reveals that, P almost surely,

$$\limsup_{t\to\infty} n^{-1} \log^+ \tilde{A}_n(y) = \begin{cases} 0 & \text{if } P(A(y)\log^+ A(y)) < \infty, \\ \infty & \text{if } P(A(y)\log^+ A(y)) = \infty. \end{cases} \tag{46}$$

Now suppose that $\lambda \in (\tilde{\lambda}, 0]$ and $P(A(y)\log^+ A(y)) = \infty$ for some $y \in I$ (with $r(y) > 0$). Let S_k be the time of the k^{th} fission along the spine producing $\tilde{A}_k(\eta_{S_k})$ additional particles, then

$$Z_\lambda(S_k) \ge \tilde{A}_k(\eta_{S_k}) v_\lambda(\eta_{S_k})\, e^{\lambda(\xi_{S_k} + c_\lambda S_k)}$$

where $(\xi_t + c_\lambda t)/t \to -\lambda c'_\lambda > 0$, η_t is ergodic so the event $\{\eta_{S_k} = y\}$ will occur for infinitely many k since $r(y) > 0$, and $n_t/t \to < Rv_\lambda, v_\lambda >_\pi$ so $S_k/k \to < Rv_\lambda, v_\lambda >_\pi^{-1}$, hence the super-exponential growth for $\tilde{A}_k(y)$ from (46) gives $\limsup_{t\to\infty} Z_\lambda(t) = \infty$ $\tilde{\mathbb{Q}}_\lambda$-almost surely which then implies that $P(Z_\lambda(\infty) = 0) = 1$.

Finally, suppose that $\lambda \in (\tilde{\lambda}, 0]$ and $P(A(y)\log^+ A(y)) < \infty$ for all $y \in I$. Recall from (26):

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{k=1}^{n_t} \tilde{A}_k(\eta_{S_k}) v_\lambda(\eta_{S_k})\, e^{\lambda(\xi_{S_k} + c_\lambda S_k)} + v_\lambda(\eta_t) e^{\lambda(\xi_t + c_\lambda t)}. \tag{47}$$

In this case, the facts that $(\xi_t + c_\lambda t)/t \to -\lambda c'_\lambda > 0$ and $S_k/k \to < Rv_\lambda, v_\lambda >_\pi^{-1}$ together with the moment conditions and (46) implying that the $\tilde{A}_k(y)$'s all have sub-exponential growth means that

$$\limsup_{t\to\infty} \tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) < \infty \qquad \tilde{\mathbb{Q}}_\lambda\text{-a.s.}$$

Fatou's lemma then gives $\liminf_{t\to\infty} Z_\lambda(t) < \infty$, $\tilde{\mathbb{Q}}_\lambda$-a.s., hence also $\mathbb{Q}_\lambda$-a.s. In addition, since $Z_\lambda(t)^{-1}$ is a positive $\mathbb{Q}_\lambda$-martingale (recall (39)) with an almost sure limit, this means that $\lim_{t\to\infty} Z_\lambda(t) < \infty$, $\mathbb{Q}_\lambda$-a.s. and then (45) yields that $P(Z_\lambda(\infty)) = 1$ and so $Z_\lambda(t)$ converges almost surely and in $\mathcal{L}^1(P)$. □

Discussion of rate of convergence to zero and left-most particle speed.

Alternatively, when $\lambda < \tilde{\lambda}$ we can readily obtain the rate of convergence to zero with the following simple argument, adapted from Git *et al.* [20]. By Proposition 9.1,

$$Z_\lambda(t)^q \le \sum_{u\in N(t)} v_\lambda(Y_u(t))^q\, e^{q\lambda(X_u(t)+c_{q\lambda})}\, e^{q\lambda(c_\lambda-c_{q\lambda})t} \le K\, Z_{q\lambda}(t) e^{q\lambda(c_\lambda-c_{q\lambda})t}$$

where $K := \max_{y\in I} v_\lambda^q(y)/v_{q\lambda}(y) < \infty$ since I is finite and $v_\lambda > 0$. Recall that c_λ has a minimum over $\lambda \in (-\infty, 0]$ at $\tilde\lambda$ with $c_{\tilde\lambda} = -E_{\tilde\lambda} = -\tilde\lambda < Av_{\tilde\lambda}, v_{\tilde\lambda} >_\pi$. Then, since $Z_{q\lambda}(t)$ is a convergent martingale, we can choose q such that $q\lambda = \tilde\lambda$ giving $Z_\lambda(t)$ decaying exponentially to zero at least at rate $\lambda(c_\lambda - c_{\tilde\lambda})$.

Further, once we know that P and $\mathbb{Q}_\lambda$ are equivalent for every $\lambda \in (\tilde\lambda, 0]$, since the spine moves such that $\xi_t/t \to -c_\lambda - \lambda c'_\lambda$ under $\mathbb{Q}_\lambda$, the left-most particle $L(t) := \inf_{u\in N(t)} X_u(t)$ must satisfy $\liminf_t L(t)/t \le -c_{\tilde\lambda}$, P-a.s. On the other hand, the convergence of the Z_λ P-martingales quickly gives the same upper bound on the fastest speed of any particle, leading to $L(t)/t \to -c_{\tilde\lambda}$, P-a.s. This result also reveals that the rate of exponential decay found above is actually best possible.

10.6 Proof of Theorem 10.5

Proof of Part 1:

Suppose $p \in (1, 2]$, then with $q := p - 1$ a slight modification of the BBM proof arrives at

$$\begin{aligned} P^{x,y}(Z_\lambda(t)^p) &= e^{\lambda x} v_\lambda(y) \tilde{\mathbb{Q}}_\lambda^{x,y}(Z_\lambda(t)^q) \\ &\le e^{\lambda x} v_\lambda(y) \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{k=1}^{n_t} A_k^q v_\lambda(\eta_{S_k})^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\Big) \\ &\quad + e^{\lambda x} v_\lambda(y) \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big) \end{aligned}$$

and the proof of $\mathcal{L}^p$-boundedness will be complete once we show that this RHS expectation is bounded in t.

The spine term. Since I is finite we note that $\langle v_\lambda^p, v_{p\lambda}\rangle_\pi < \infty$. It is always useful to first focus on the spine term, since we can change the measure with (41) to get

$$\begin{aligned} \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big) &= \tilde{P}^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}. \frac{v_\lambda(\eta_t)\, e^{\int_0^t MR(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}}{v_\lambda(y) e^{\lambda x}}\Big) \\ &= e^{q\lambda x} \frac{v_{p\lambda}(y)}{v_\lambda(y)} g_t(y) e^{-(pE_\lambda - E_{p\lambda})t} \end{aligned} \tag{48}$$

where, for all $y \in I$

$$g_t(y) := \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big) \to \langle v_\lambda^p, v_{p\lambda}\rangle_\pi$$

as $t \to \infty$ and $\langle g_t v_{p\lambda}, v_{p\lambda}\rangle_\pi = \langle v_\lambda^p, v_{p\lambda}\rangle_\pi$ for all $t \ge 0$, since η_t is a finite-state irreducible Markov chain under $\tilde{\mathbb{Q}}_\mu$ with invariant distribution $\pi_\mu(y) = v_\mu(y)^2\pi(y)$. It follows that the long term the growth or decay of the spine term is determined by the sign of $pE_\lambda - E_{p\lambda}$.

The sum term. We now assume that $pE_\lambda - E_{p\lambda} > 0$. We know that under $\tilde{\mathbb{Q}}_\lambda$ and conditional on knowing η, the fission times $\{S_k : k \ge 0\}$ on the spine occur as a

Poisson process of rate $(1+m(\eta_s))r(\eta_s)$ with the k^{th} fission yielding an additional A_k offspring, each A_k being an independent copy of $\tilde{A}(y)$ which has the size-biased distribution $\{(1+k)p_k(y)/(1+m(y)) : k \geq 0\}$ where $y = \eta_{S_k}$ is the type at the time of fission. We also recall from Lemma 9.2 that

$$M_q(y) := \tilde{\mathbb{Q}}_\lambda(\tilde{A}^q(y)) < \infty \iff P(A^p(y)) < \infty.$$

Therefore, if we condition on $\mathcal{G}_t$ which knows about (ξ_s, η_s) at all times $0 \leq s \leq t$ we can transform the sum into an integral, use Fubini's theorem and the change of measure used in (48):

$$\begin{aligned}
&\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{k=1}^{n_t} A_k^q v_\lambda(\eta_{S_k})^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\Big) \\
&\qquad = \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{k=1}^{n_t} A_k^q v_\lambda(\eta_{S_k})^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k} \big| \mathcal{G}_t\Big)\Big) \\
&\qquad = \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\int_0^t (1+m(\eta_s))r(\eta_s)\, M_q(\eta_s) v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\,\mathrm{d}s\Big) \\
&\qquad = \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big((1+m(\eta_s))r(\eta_s) M_q(\eta_s) v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\Big)\,\mathrm{d}s \\
&\qquad = e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \int_0^t h_s(\eta_s) e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s \\
&\qquad = e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \frac{k_t(y)}{pE_\lambda - E_{p\lambda}}
\end{aligned}$$

where

$$h_s(y) := \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\tilde{r}(\eta_s) M_q(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big), \qquad \tilde{r}(y) := (1+m(y))r(y),$$

$$\text{and} \qquad k_t(y) := \mathbb{E}(h_U(y); U \leq t)$$

with U an independent exponential of rate $(pE_\lambda - E_{p\lambda}) > 0$. Note that, for all $y \in I$, $h_s(y) \to \langle \tilde{r} M_q v_\lambda^p, v_{p\lambda}\rangle$ and $k_t(y) \uparrow k_\infty(y)$ as $t \to \infty$, where $\langle k_t v_{p\lambda}, v_{p\lambda}\rangle = \langle \tilde{r} M_q v_\lambda^p, v_{p\lambda}\rangle \mathbb{P}(U \leq t) \uparrow \langle k_\infty v_{p\lambda}, v_{p\lambda}\rangle = \langle \tilde{r} M_q v_\lambda^p, v_{p\lambda}\rangle$. Then, since $M_q(w) < \infty \iff P(A(w)^p) < \infty$, and I is finite, we are guaranteed that $k_\infty(y) < \infty$ for all $y \in I$ as long as $P(A(w)^p) < \infty$ for all $w \in I$.

Having dealt with both the spine term and the sum term, we have obtained the upper-bound

$$P^{x,y}\big(Z_\lambda(t)^p\big) \leq \frac{e^{p\lambda x} v_{p\lambda}(y)}{(pE_\lambda - E_{p\lambda})}\Big(k_t(y) + g_t(y)\,(pE_\lambda - E_{p\lambda})\,e^{-(pE_\lambda - E_{p\lambda})t}\Big)$$

and since $Z_\lambda(t)^p$ is a P-submartingale, we find that

$$P^{x,y}\big(Z_\lambda(t)^p\big) \leq \frac{e^{p\lambda x} v_{p\lambda}(y)}{(pE_\lambda - E_{p\lambda})}\,k_\infty(y) \qquad (\forall t \geq 0)$$

and $Z_\lambda(t)$ will be bounded in $\mathcal{L}^p(P^{x,y})$ if we have both $pE_\lambda - E_{p\lambda} > 0$ and $P(A^p(w)) < \infty$ for all $w \in I$. $\square$

Proof of Part 2:

The earlier proof for BBM goes through with minor modification. Exactly as in the BBM case, looking only at the contribution of the spine means that Z_λ is unbounded in $\mathcal{L}^p(P^{x,y})$ if $pE_\lambda - E_{p\lambda} < 0$. In addition, letting T be any fission time along the path of the spine,

$$Z_\lambda(T) \geq (1 + \tilde{A}(\eta_T))v_\lambda(\eta_T)e^{\lambda\xi_T - E_\lambda T}$$

where $\tilde{A}(\eta_T)$ is the number of additional offspring produced at the time of fission. Then, with $m_q(y) := \tilde{\mathbb{Q}}((1+\tilde{A}(y))^q) < \infty \iff P(A^p(y)) < \infty$,

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}(Z_\lambda(T)^q) &\geq e^{q\lambda x}\,\tilde{\mathbb{Q}}_\lambda^{x,y}(m_q(\eta_T)v_\lambda(\eta_T)^q e^{q\lambda\xi_T - qE_\lambda T}) \\ &= e^{q\lambda x}\,\tilde{\mathbb{Q}}_{p\lambda}^{x,y}\left(m_q(\eta_T)\frac{v_\lambda(\eta_T)^p}{v_{p\lambda}(\eta_T)}\,e^{-(pE_\lambda - E_{p\lambda})T}\right)\end{aligned}$$

and so Z_λ will also be unbounded in $\mathcal{L}^p(P^{x,y})$ if $m_q(y) = \infty \iff P(A^p(y)) = \infty$ for any $y \in I$ (taking a fission time when also in state y). □

Remarks on signed martingales and Kesten-Stigum type theorems

In the multi-typed BBM, for each λ there will be other (signed) additive martingales corresponding to the different eigenvectors and eigenvalues obtained from solving (34); the Z_λ martingale simply corresponds to the Perron-Frobenius, or ground-state, eigenvalue E_λ and (strictly positive) eigenvector v_λ. Since $|u+v|^q \leq (|u| + |v|)^q \leq |u|^q + |v|^q$ for all $u, v \in \mathbb{R}$, the above proof will also adapt to give convergence results for signed martingales. In fact, when there is a complete orthonormal set of eigenvectors, a Kesten-Stigum like theorem would then swiftly follow (for example, see Harris [30] in the context of the continuous-type model of the next section).

11 A continuous-typed branching-diffusion

The previous finite-type model was originally inspired by the model that we now turn to, originally laid out in Harris and Williams [28]. In this model the type moves on the real line as an Orstein-Uhlenbeck process associated with the generator

$$Q_\theta := \frac{\theta}{2}\Big(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\Big), \qquad \text{with } \theta > 0 \text{ considered as the } \textit{temperature},$$

which has the standard normal density as its invariant distribution:

$$\pi(y) := (2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}y^2}.$$

The spatial movement of a particle of type y is a driftless Brownian motion with instantaneous variance

$$A(y) := ay^2, \qquad \text{for some fixed } a > 0,$$

and fission of a particle of type y occurs at a rate

$$R(y) := ry^2 + \rho, \qquad \text{where } r, \rho > 0 \text{ are fixed},$$

to produce two particles at the same type-space location as the parent (we consider only binary splitting). The model has very different behaviour for low temperature values (i.e. low θ), but most studies have considered the high temperature regime where $\theta > 8r$. Also, the parameter λ must be restricted to an interval $(\lambda_{\min}, 0)$ in order for some of the model's parameters to remain in $\mathbb{R}$, where

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}}.$$

Generally, *unboundedness* in a model's rates is a serious obstacle to classical proofs since they often depend on the expectation semigroup of the branching process, and unbounded rates tend to lead to unbounded *eigenfunctions*. Here this is the case, but the existence of a spectral theory for their particular expectation operator allowed Harris and Williams to get a sufficiently good bound in particular for a *non-linear* term (see Theorem 5.1 of [28]), and therefore to prove $\mathcal{L}^p$-convergence of the martingale. Other convergence results for various martingales and weighted sums over particles for this model also appear in Harris [30], again using more classical methods and requiring 'non-linear' calculations. The spine approach we again adopt here is both simple and more generic in nature; requiring no such special 'non-linear' calculations, it elegantly produces very good estimates that only involve easy one-particle calculations.

We use the same notation as previously, $\mathbb{X}_t = \{(X_u(t), Y_u(t)) : u \in N_t\}$ to denote the point process of space-type locations in $\mathbb{R} \times \mathbb{R}$, and suppose that the measures $\{\tilde{P}^{x,y} : (x, y) \in \mathbb{R}^2\}$ on the natural filtration with a spine $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ are such that the initial ancestor starts at (x, y) and $(\mathbb{X}_t, (\xi_t, \eta_t))$ becomes the above-described branching diffusion with a spine.

11.1 The measure change

Although there are some significant differences, this model is similar in flavour to our finite-type model. There is a strictly-positive martingale Z_λ defined as

$$Z_\lambda(t) := \sum_{u \in N_t} v_\lambda(Y_u(t)) e^{\lambda X_u(t) - E_\lambda t}$$

where v_λ and E_λ are the eigenvector and eigenvalue associated with the self-adjoint (in $\mathcal{L}^2(\pi)$) operator:

$$Q_\theta + \frac{1}{2}\lambda^2 A(y) + R(y).$$

The eigenfunction v_λ is normalizable against the $\mathcal{L}^2(\pi)$ norm, and can be found explicitly as

$$v_\lambda(y) = e^{\psi_\lambda^- y^2}$$

where

$$\psi_\lambda^- := \frac{1}{4} - \frac{\mu_\lambda}{2\theta}, \quad \mu_\lambda := \frac{1}{2}\sqrt{\theta^2 - \theta(8r + 4a\lambda^2)},$$

are both positive for all $\lambda \in (\lambda_{\min}, 0)$; another important parameter is $\psi_\lambda^+ := \frac{1}{4} + \frac{\mu_\lambda}{2\theta}$. The eigenvalue E_λ is then given by

$$E_\lambda = \rho + \theta\psi_\lambda^-.$$

We again define the speed function $c_\lambda := -E_\lambda/\lambda$, and $\tilde{\lambda}(\theta) < 0$ is the unique point (on the negative axis) at which c_λ hits its minimum $\tilde{c}(\theta)$ – further details are given in Harris and Williams [28]. We are going to use spines to prove the following result, in which the critical case of $\lambda = \tilde{\lambda}$ and the necessary conditions for $\mathcal{L}^p(P)$-convergence are new results:

Theorem 11.1 *Suppose that $\lambda \in (\lambda_{\min}, 0)$.*

1. *Let $p \in (1, 2]$. The martingale Z_λ is $\mathcal{L}^p(P)$-bounded if both $pE_\lambda - E_{p\lambda} > 0$ and $p\psi_\lambda^- < \psi_{p\lambda}^+$. In particular, for all $\lambda \in (\tilde{\lambda}(\theta), 0]$, Z_λ is a uniformly-integrable martingale.*
2. *Z_λ is unbounded in $\mathcal{L}^p(P)$ if either $pE_\lambda - E_{p\lambda} < 0$ or $p\psi_\lambda^- > \psi_{p\lambda}^+$.*
3. *Almost surely under P, $Z_\lambda(\infty) = 0$ if $\lambda \le \tilde{\lambda}(\theta)$.*

Once again, for each $\lambda \le 0$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}} 2^{n_t} v_\lambda(\eta_t) e^{\lambda \xi_t - E_\lambda^- t}, \tag{49}$$

so that with $\mathbb{Q}_\lambda := \tilde{\mathbb{Q}}_\lambda|_{\mathcal{F}_\infty}$ we have

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

The facts are that under $\tilde{\mathbb{Q}}_\lambda$:

- the spine diffusion ξ_t has instantaneous drift $a\eta_t^2\lambda$;
- the type process η_t has generator $\frac{\theta}{2}\left(\frac{\partial^2}{\partial y^2} - \frac{2\mu_\lambda}{\theta} y \frac{\partial}{\partial y}\right)$ and an invariant probability measure $\pi_\lambda := \langle v_\lambda, v_\lambda \rangle_\pi^{-1} v_\lambda^2 \pi$, corresponding to a normal distribution, $N(0, \frac{\theta}{2\mu_\lambda})$;
- fission times on the spine occur at the accelerated rate of $2R(\eta_t)$;
- all particles not in the spine behave as if under the original measure P.

We briefly comment that, along similar lines as discussed for the finite-typed BBM case, we could now give a straightforward spine proof that the asymptotic right-most particle speed in this continuous typed BBM model is almost surely $\tilde{c}(\theta)$.

11.2 Proof of Theorem 11.1

Proof of Part 1: Suppose $p \in (1, 2]$. Then using the spine decomposition with Jensen's inequality and Proposition 9.1 we find,

$$P^{x,y}(Z_\lambda^-(t)^p) \le e^{\lambda x} v_\lambda(y) \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big)$$
$$+ e^{\lambda x} v_\lambda(y) \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big).$$

Assume that $pE_\lambda - E_{p\lambda} > 0$ and $p\psi_\lambda^- < \psi_{p\lambda}^+$. As seen in Harris and Williams [28], we can do many calculations explicitly in this model, largely due to the fact that under $\tilde{\mathbb{Q}}_{p\lambda}^{0,y}$

$$\eta_s \sim \mathrm{N}\left(e^{-\mu_{p\lambda}s}y, \frac{\theta(1-e^{-2\mu_{p\lambda}s})}{2\mu_{p\lambda}}\right) \to \mathrm{N}\left(0, \frac{\theta}{2\mu_{p\lambda}}\right)$$

and the eigenfunctions v_λ have such simple exponential form. For example,

$$\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\left(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\right) = \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\left(e^{(p\psi_\lambda^- - \psi_{p\lambda}^-)\eta_s^2}\right) \tag{50}$$

can easily be seen to be finite and bounded for all $s \geq 0$ if and only if $p\psi_\lambda^- - \psi_{p\lambda}^- - \frac{\mu_{p\lambda}}{\theta} = p\psi_\lambda^- - \psi_{p\lambda}^+ < 0$, and just as readily calculated explicitly.

In fact, more 'natural' conditions for $\mathcal{L}^p$ convergence of the martingales would be that

$$\left\langle RM_q v_\lambda^p, v_{p\lambda}\right\rangle_\pi < \infty, \qquad \left\langle v_\lambda^p, v_{p\lambda}\right\rangle_\pi < \infty, \qquad \text{and} \qquad pE_\lambda - E_{p\lambda} < 0,$$

where $M_q(y) := \tilde{\mathbb{Q}}(\tilde{A}^q(y))$ with $\tilde{A}$ the size-biased offspring distribution (here, binary splitting means $\tilde{A}(y) \equiv 1$), and we present arguments below that are more generic in nature, at least in terms of adapting to other 'suitably' ergodic type motions and random family sizes. Note, the last condition above is related to the natural convexity of E_λ and, in our specific model, both integrability conditions are guaranteed by $p\psi_\lambda^- - \psi_{p\lambda}^+ < 0$.

The spine term. On the algebra $\mathcal{G}_t$ the change of measure takes the form

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\mathcal{G}_t} = \frac{v_\lambda(\eta_s)}{v_\lambda(y)}\exp\left(\int_0^t R(\eta_s)\,\mathrm{d}s + \lambda(\xi_t - x) - E_\lambda^- t\right),$$

which we can use on the spine term to arrive at

$$f_t(x,y) := e^{\lambda x}v_\lambda(y)\tilde{\mathbb{Q}}_\lambda^{x,y}\left(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\right) = e^{p\lambda x}v_{p\lambda}(y)\,g_t(y)e^{-(pE_\lambda - E_{p\lambda})t} \tag{51}$$

with $g_t(y) := \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\left(v_\lambda^p(\eta_t)/v_{p\lambda}(\eta_t)\right)$. Under the assumption that $p\psi_\lambda^- < \psi_{p\lambda}^+$, it easy to check that $\left\langle v_\lambda^p, v_{p\lambda}\right\rangle_\pi < \infty$, that is $v_\lambda^p/v_{p\lambda} \in L^1(\pi_{p\lambda})$ from which it follows that $g_t \in L^1(\pi_{p\lambda})$ for all $t \geq 0$. Since η has equilibrium $\pi_{p\lambda}$ under $\tilde{\mathbb{Q}}_{p\lambda}$, we find $\left\langle g_t v_{p\lambda}, v_{p\lambda}\right\rangle_\pi = \left\langle v_\lambda^p, v_{p\lambda}\right\rangle_\pi < \infty$ and $g_t(y) \to \left\langle v_\lambda^p, v_{p\lambda}\right\rangle_\pi \left\langle v_{p\lambda}, v_{p\lambda}\right\rangle_\pi^{-1} < \infty$ as $t \to \infty$ for all $y \in \mathbb{R}$.

We also note that since $g_t \in L^1(\pi_{p\lambda})$, we have $f_t \in L^1(\tilde{\pi}_{p\lambda})$ where $\tilde{\pi}_\mu := \left\langle 1, v_\mu\right\rangle_\pi^{-1} v_\mu \pi$ and then

$$\int_{y\in\mathbb{R}} \tilde{\pi}_{p\lambda}(y) f_t(x,y)dy = e^{p\lambda x}\frac{\left\langle v_\lambda^p, v_{p\lambda}\right\rangle_\pi}{\left\langle 1, v_{p\lambda}\right\rangle_\pi}\, e^{-(pE_\lambda - E_{p\lambda})t}.$$

The sum term. Note that under the parameter assumptions we have $\left\langle Rv_\lambda^p, v_{p\lambda}\right\rangle_\pi < \infty$. As for the finite-type model the fission times S_u on the spine occur as a Cox process and therefore

$$\begin{aligned} g_t(x,y) &:= e^{\lambda x} v_\lambda(y) \tilde{\mathbb{Q}}_\lambda^{x,y} \Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda \xi_{S_u} - qE_\lambda S_u} \Big) \\ &= e^{\lambda x} v_\lambda(y) \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y} \Big(2R(\eta_s)\, v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s} \Big)\, \mathrm{d}s \\ &= e^{\lambda x} v_\lambda(y) \int_0^t \tilde{\mathbb{Q}}_{p\lambda}^{0,y} \Big(2R(\eta_s)\, \frac{v_\lambda^p}{v_{p\lambda}}(\eta_s) \Big) e^{-(pE_\lambda - E_{p\lambda})s}\, \mathrm{d}s \\ &= e^{p\lambda x} v_{p\lambda}(y) k_t(y) \end{aligned}$$

where

$$k_t(y) := \int_0^t h_s(y) e^{-(pE_\lambda - E_{p\lambda})s}\, \mathrm{d}s, \qquad h_s(y) := \tilde{\mathbb{Q}}_{p\lambda}^{0,y} \Big(2R(\eta_s)\, \frac{v_\lambda^p}{v_{p\lambda}}(\eta_s) \Big)$$

and $h_t, k_t \in L^1(\pi_{p\lambda})$. Note, $k_t(y) \uparrow k_\infty(y) \in L^1(\pi_{p\lambda})$ as $t \to \infty$ where

$$\begin{aligned} \frac{\langle k_t v_{p\lambda}, v_{p\lambda} \rangle_\pi}{\langle v_{p\lambda}, v_{p\lambda} \rangle_\pi} &= \langle 2R v_\lambda^p, v_{p\lambda} \rangle_\pi \frac{\Big(1 - e^{-(pE_\lambda - E_{p\lambda})t}\Big)}{(pE_\lambda - E_{p\lambda})} \\ &\uparrow \frac{\langle 2R v_\lambda^p, v_{p\lambda} \rangle_\pi}{(pE_\lambda - E_{p\lambda})} = \langle k_\infty v_{p\lambda}, v_{p\lambda} \rangle_\pi < \infty. \end{aligned}$$

Note, $k_t \in L^1(\pi_{p\lambda})$ implies $g_t \in L^1(\tilde{\pi}_{p\lambda})$, with an explicit calculation again possible.

Bringing together the results for the sum and spine terms, we have an upper bound

$$P^{x,y}(Z_\lambda(t)^p) \le e^{p\lambda x} v_{p\lambda}(y) \Big\{ g_t(y) e^{-(pE_\lambda - E_{p\lambda})t} + k_t(y) \Big\} \in L^1(\tilde{\pi}_{p\lambda}) \qquad (52)$$

and hence the submartingale property reveals that

$$P^{x,y}(Z_\lambda(t)^p) \le e^{p\lambda x} v_{p\lambda}(y) k_\infty(y) < \infty$$

for all $t \ge 0$ and all $y \in \mathbb{R}$. □

Proof of Part 2: We need only dominate the martingale by the spine at time t, yielding

$$\begin{aligned} \tilde{\mathbb{Q}}_\lambda^{x,y}(Z_\lambda(t)^q) &\ge \tilde{\mathbb{Q}}_\lambda^{x,y}(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}) \\ &= e^{q\lambda x} \frac{v_{p\lambda}(y)}{v_\lambda(y)} \tilde{\mathbb{Q}}_{p\lambda}^y \left(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t) \right) e^{-(pE_\lambda - E_{p\lambda})t}. \end{aligned}$$

Hence Z_λ is unbounded in $\mathcal{L}^p(P^x)$ if either $pE_\lambda - E_{p\lambda} < 0$ or $\langle v_\lambda^p, v_{p\lambda} \rangle_\pi = \infty$. □

Proof of Part 3: The proof that we have seen in the finite-type model will work here with little change: under $\tilde{\mathbb{Q}}_\lambda$ the spatial motion is

$$\xi_t = B\left(\int_0^t a(\eta_s) \mathrm{d}s \right) + \lambda \int_0^t a(\eta_s) \mathrm{d}s,$$

and the type process η_s has invariant distribution $N(0, \frac{\theta}{2\mu_\lambda})$, whence $t^{-1}\xi_t \to \lambda a\theta/\mu_\lambda = E'_\lambda$. Therefore it follows that under $\tilde{\mathbb{Q}}_\lambda$ the diffusion $\xi_t + c_\lambda t$ drifts off to $-\infty$ if $\lambda < \tilde{\lambda}(\theta)$. When $\lambda = \tilde{\lambda}$, it is also simple to check that $\xi_t + c_\lambda t$ is recurrent, so has $\liminf\{\xi_t + c_\lambda t\} = -\infty$. Whence, in either case, bounding Z_λ below by the spine's contribution as done before, we have $Z_\lambda(t) \geq v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)}$ and since $v_\lambda > 0$ and η_t recurrent, we see that $\limsup_{t\to\infty} Z_\lambda(t) = \infty$ almost surely under $\tilde{\mathbb{Q}}_\lambda^{x,y}$. □

Acknowledgement. We would like to thank the referee for their careful reading of the original articles and for their helpful suggestions.

References

1. Søren Asmussen and Heinrich Hering, *Strong limit theorems for general supercritical branching processes with applications to branching diffusions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36** (1976), no. 3, 195–212.
2. Krishna B. Athreya, *Change of measures for Markov chains and the $L \log L$ theorem for branching processes*, Bernoulli **6** (2000), no. 2, 323–338.5 MR 2001g:60202
3. Jean Bertoin, Random fragmentation and coagulation processes, Cambridge University Press, 2006, ISBN: 9780521867283
4. J. D. Biggins and A. E. Kyprianou, *Measure change in multitype branching*, Adv. in Appl. Probab. **36** (2004), no. 2, 544–581.
5. A. Champneys, S. Harris, J. Toland, J. Warren, and D. Williams, *Algebra, analysis and probability for a coupled system of reaction-diffusion equations*, Philosophical Transactions of the Royal Society of London **350** (1995), 69–112.
6. B. Chauvin, *Arbres et processus de Bellman-Harris*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 209–232.
7. B. Chauvin and A. Rouault, *Boltzmann-Gibbs weights in the branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 41–50.
8. Brigitte Chauvin, *Product martingales and stopping lines for branching Brownian motion*, Ann. Probab. **19** (1991), no. 3, 1195–1205.
9. Brigitte Chauvin and Alain Rouault, *KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees*, Probab. Theory Related Fields **80** (1988), no. 2, 299–314.
10. Brigitte Chauvin, Alain Rouault, and Anton Wakolbinger, *Growing conditioned trees*, Stochastic Process. Appl. **39** (1991), no. 1, 117–130.
11. Richard Durrett, *Probability: Theory and examples*, 2nd ed., Duxbury Press, 1996.
12. János Engländer and Andreas E. Kyprianou, *Local extinction versus local exponential growth for spatial branching processes*, Ann. Probab. **32** (2004), no. 1A, 78–99.
13. János Engländer, *Branching diffusions, superdiffusions and random media*, Probab. Surveys Vol. **4** (2007) 303-364.
14. János Engländer,, Simon C. Harris andAndreas E. Kyprianou, *Laws of Large numbers for spatial branching processes*, Annales de l'Institut Henri Poincar (B) Probability and Statistics, (2009), to appear.

15. Jochen Geiger, *Size-biased and conditioned random splitting trees*, Stochastic Process. Appl. **65** (1996), no. 2, 187–207.
16. Jochen Geiger, *Elementary new proofs of classical limit theorems for Galton-Watson processes*, J. Appl. Probab. **36** (1999), no. 2, 301–309.
17. Jochen Geiger, *Poisson point process limits in size-biased Galton-Watson trees*, Electron. J. Probab. **5** (2000), no. 17, 12 pp. (electronic).
18. Jochen Geiger and Lars Kauffmann, *The shape of large Galton-Watson trees with possibly infinite variance*, Random Structures Algorithms **25** (2004), no. 3, 311–335.
19. Hans-Otto Georgii and Ellen Baake, *Supercritical multitype branching processes: the ancestral types of typical individuals*, Adv. in Appl. Probab. **35** (2003), no. 4, 1090–1110.
20. Y. Git, J. W. Harris, and S. C. Harris, *Exponential growth rates in a typed branching diffusion*, Ann. App. Probab., **17** (2007), no. 2, 609-653. `doi:10.1214/105051606000000853`
21. Robert Hardy, *Branching diffusions*, Ph.D. thesis, University of Bath Department of Mathematical Sciences, 2004.
22. Robert Hardy and Simon C. Harris, *Some path large deviation results for a branching diffusion*, (2007), submitted.
23. Robert Hardy and Simon C. Harris, *A conceptual approach to a path result for branching Brownian motion*, Stoch. Proc. and Applic., **116** (2006), no. 12, 1992–2013. `doi:10.1016/j.spa.2006.05.010`
24. Robert Hardy and Simon C. Harris, *A new formulation of the spine approach for branching diffusions*, (2006). `arXiv:math.PR/0611054`
25. Robert Hardy and Simon C. Harris, *Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales*, (2006). `arXiv:math.PR/0611056`
26. J. W. Harris, S. C. Harris, and A. E. Kyprianou, *Further probabilistic analysis of the Fisher-Kolmogorov-Petrovskii-Piscounov equation: one-sided travelling waves*, Ann. Inst. H. Poincaré Probab. Statist. **42** (2006), no. 1, 125–145.
27. John W. Harris and Simon C. Harris, *Branching Brownian motion with an inhomogeneous breeding potential*, Ann. Inst. H. Poincaré Probab. Statist. (2008), to appear.
28. S. C. Harris and D. Williams, *Large deviations and martingales for a typed branching diffusion. I*, Astérisque (1996), no. 236, 133–154, Hommage à P. A. Meyer et J. Neveu.
29. Simon C. Harris, *Travelling-waves for the FKPP equation via probabilistic arguments*, Proc. Roy. Soc. Edinburgh Sect. A **129** (1999), no. 3, 503–517.
30. Simon C. Harris, *Convergence of a "Gibbs-Boltzmann" random measure for a typed branching diffusion*, Séminaire de Probabilités, XXXIV, Lecture Notes in Math., vol. 1729, Springer, Berlin, 2000, pp. 239–256.
31. S.C. Harris, R. Knobloch and A.E. Kyprianou *Strong Law of Large Numbers for Fragmentation Processes*, (2008) arXiv:0809.2958v1, submitted.
32. Theodore E. Harris, *The theory of branching processes*, Dover ed., Dover, 1989.
33. Yueyun Hu and Zhan Shi, *Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees*, (2008) Ann. App. Probab., to appear.
34. Aleksander M. Iksanov, *Elementary fixed points of the BRW smoothing transforms with infinite number of summands*, Stochastic Process. Appl. **114** (2004), no. 1, 27–50.

35. O. Kallenberg, *Foundations of modern probability*, Springer-Verlag, 2002.
36. H. Kesten and B.P. Stigum, *Additional limit theorems for indecomposable multi-dimensional Galton-Watson processes*, Ann. Math. Stat. **37** (1966), 1463–1481.
37. H. Kesten and B.P. Stigum, *A limit theorem for multidimensional Galton-Watson processes*, Ann. Math. Stat. **37** (1966), 1211–1223.
38. H. Kesten and B.P. Stigum, *Limit theorems for decomposable multi-dimensional Galton-Watson processes*, J. Math. Anal. Applic. **17** (1967), 309–338.
39. Erwin Kreyszig, *Introductory functional analysis with applications*, John Wiley and Sons, 1989.
40. A. E. Kyprianou, *Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris' probabilistic analysis*, Ann. Inst. H. Poincaré Probab. Statist. **40** (2004), no. 1, 53–72.
41. Thomas Kurtz, Russell Lyons, Robin Pemantle, and Yuval Peres, *A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 181–185.
42. Quansheng Liu and Alain Rouault, *On two measures defined on the boundary of a branching tree*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 187–201.
43. Russell Lyons, *A simple path to Biggins' martingale convergence for branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221.
44. Russell Lyons, Robin Pemantle, and Yuval Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, Ann. Probab. **23** (1995), no. 3, 1125–1138.
45. J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207.
46. J. Neveu, *Multiplicative martingales for spatial branching processes*, Seminar on Stochastic Processes (E. Çinlar, K.L.Chung, and R.K.Getoor, eds.), Birkhäuser, 1987, pp. 223–241.
47. Peter Olofsson, *The $x \log x$ condition for general branching processes*, J. Appl. Probab. **35** (1998), no. 3, 537–544.
48. E. Seneta, *Non-negative matrices and Markov chains*, Springer-Verlag, 1981.
49. Edward C. Waymire and Stanley C. Williams, *A general decomposition theory for random cascades*, Bull. Amer. Math. Soc. (N.S.) **31** (1994), no. 2, 216–222.